Parameter Redundancy with Applications in Statistical Ecology

Ben Arthur Hubbard

A Thesis submitted for the degree of Doctor

of Philosophy in the subject of Statistics

School of Mathematics, Statistics and Actuarial Science

University of Kent

December 2014

Abstract

This thesis is concerned with parameter redundancy in statistical ecology models. If it is not possible to estimate all the parameters, a model is termed parameter redundant. Parameter redundancy commonly occurs when parameters are confounded in the model so that the model could be reparameterised in terms of a smaller number of parameters. In principle, it is possible to use symbolic algebra to determine whether or not all the parameters of a certain ecological model can be estimated using classical methods of statistical inference.

We examine a variety of different ecological models: We begin by exploring models based on marking a number of animals and observing the same animals at future time points. These observations can either be when the animal is marked and then recovered dead in mark-recovery modelling, or when the animal is marked and then recaptured alive in capture-recapture modelling. We also explore capture-recapture-recovery models where both dead recoveries and alive recaptures can be observed in the same study. We go on to explore occupancy models which are used to obtain estimates of the probability of presence, or absence, for living species by the use of repeated detection surveys, where these models have the advantage that individuals are not required to be marked. A variety of different occupancy models are examined included the addition of season-dependent parameters, group-dependent parameters and species-dependent, along with other models.

We investigate parameter redundancy by deriving general results for a variety of different models where the model's parameter dependencies can be relaxed suited to different studies. We also analyse how the results change for specific data sets and how sparse data influence whether or not a model is parameter redundant using procedures written in `Maple`. This theory on parameter redundancy is vital for the correct use of these ecological models so that valid statistical inference can be made.

Acknowledgements

I would like to express my gratitude to my academic supervisors, Dr. Diana Cole and Prof. Byron Morgan, for their continued support during this PhD process and assisting me during these first steps of academic research! They have enabled me to work as an independent researcher with their helpful advice, particularly for the written elements of the thesis. I thank them for their support during the publication of the papers we have worked on and for their advice to help me present my work to a wider audience. I would also like to thank Prof. Martin Ridout from the University of Kent, and Dr. Larissa Bailey from Colorado State University, USA, for being examiners for my thesis. I would further like to recognise the School of Mathematics, Statistics and Actuarial Science at the University of Kent for their commitment to me during my eight years of undergraduate and postgraduate studies and all the staff I have been lectured and assisted by.

I am extremely grateful to the Engineering and Physical Sciences Research Council (EPSRC) and the School of Mathematics, Statistics and Actuarial Science for the financial support of my studies; without them this research would have not be possible. I acknowledge the work done by the National Centre for Statistical Ecology (NCSE) which has allowed me to present my work to colleagues and allowed provoking discussion about my research topics to be had.

I finally, but not least importantly, would like to acknowledge how amazing my parents, Leslie and Linda, have been to me and thank all their support to me, financial or otherwise. I would also to thank all my friends during my studies, such as the members of the University of Kent Snooker & Pool Club, as well as my partner Hannah for her continued support and encouragement over the years. I feel extremely privileged to be able to do this research and would like to thank all of my family and friends during this long journey!

Contents

Abstract i

Acknowledgements ii

List of Supplementary Maple Files x

1 Introduction 1

1.1 Introduction to parameter redundancy 2

1.2 Thesis structure 5

2 Mark-Recovery Models with Parameter Redundancy Examples 10

2.1 Mark-recovery models 10

2.2 What is parameter redundancy and how can it be detected? 15

2.3 Methods for finding general parameter redundancy results 23

2.4 Choices of exhaustive summary 31

2.5 Non-symbolic methods for obtaining parameter redundancy 35

2.6 Intrinsic and extrinsic parameter redundancy 39

2.7 The PLUR decomposition and near parameter redundancy 41

2.8 Detecting parameter redundancy in models with covariates 45

2.9 Intrinsic mark-recovery model results 47

2.10 Intrinsic mark-recovery models with juvenile survival probabilities . . . 49

2.11 Discussion 54

3 Capture-Recapture Models 60

3.1 Capture-recapture background 60

3.2 Capture-recapture models for m-array data 63

3.3 A transformation between capture-recapture models for m-array data and mark-recovery models for d-array data 68

3.4 The capture-history formulation of capture-recapture models and a simpler exhaustive summary 69

3.5 Intrinsic capture-recapture y/z model results 76
3.6 Capture-recapture models with juvenile survival probabilities 77
3.7 Extrinsic parameter redundancy: The dippers data set 84
3.8 An extrinsic sparseness statistic for capture-recapture models 90
3.9 Discussion 92

4 Capture-Recapture-Recovery Models **95**
4.1 Capture-recapture-recovery background 95
4.2 Life-histories and how they can form an exhaustive summary 97
4.3 Links to other capture-recapture-recovery likelihoods 100
4.4 A simpler exhaustive summary for capture-recapture-recovery models . 103
4.5 Intrinsic parameter redundancy results 109
4.6 Extrinsic parameter redundancy and the cormorants data set 114
4.7 An extrinsic sparseness statistic for capture-recapture-recovery models . 124
4.8 Fidelity in the capture-recapture-recovery model 125
4.9 Intrinsic parameter redundancy for the fidelity model 135
4.10 Extrinsic parameter redundancy for the fidelity model 136
4.11 Discussion 140

5 Occupancy Models **144**
5.1 Occupancy modelling background 144
5.2 The basic occupancy model 146
5.3 Adding more detection surveys 147
5.4 Increasing the number of seasons 149
5.5 Intrinsic parameter redundancy for occupancy models 154
5.6 Extrinsic parameter redundancy for occupancy models 155
5.7 Addition of group-dependent parameters 159
5.8 Missing observations in occupancy models 166
5.9 The effect of data sparseness on parameter redundancy 172
5.10 Discussion 178

6 Complex Occupancy Models **179**
6.1 The multiple-states occupancy model 181
6.1.1 A three-state occupancy model 181
6.1.2 A general m-state model which involves group-dependence 187
6.1.3 A general group-dependent m-state model which contains missing observations 190

6.1.4 Extrinsic parameter redundancy and Maple code for the multiple-states model . . . 192
6.1.5 A real data example: Green frogs in Maryland . . . 197
6.2 Including a species original occupancy state . . . 200
6.3 The two-species interaction occupancy model . . . 203
6.3.1 Model formulation . . . 204
6.3.2 Having more than two species in the model . . . 209
6.3.3 Parameter redundancy in the two-species model . . . 211
6.3.4 A real data example: Different species in the Yellowstone and Grand Teton data set . . . 213
6.4 The Royle-Nichols model . . . 216
6.4.1 Obtaining abundance estimates from occupancy models . . . 216
6.4.2 Adding group- and/or survey-dependency in the Royle-Nichols Poisson model . . . 218
6.4.3 Extrinsic parameter redundancy in the Royle-Nichols Poisson model . . . 220
6.5 Discussion and further occupancy models . . . 224

7 Conclusions and Future Work **227**
7.1 Discussion . . . 227
7.2 Further work . . . 229

A Common Maple Procedures **231**

B Reparameterisation Proofs **235**
B.1 Proof of the multiple capture-recapture simpler exhaustive summary theorem . . . 235
B.2 Proof of the multiple capture-recapture-recovery simpler exhaustive summary theorem . . . 242
B.3 Proof of the intrinsic fidelity model results . . . 250
B.4 Proof of the intrinsic occupancy model theorem . . . 260

List of Tables

1.1a The different ecological models that have been examined for parameter redundancy, up to December 2014 (Table 1) 6

1.1b The different ecological models that have been examined for parameter redundancy, up to December 2014 (Table 2) 7

2.1 A d-array representation of data from a mark-recovery study of lapwings from Catchpole et al. (1999) 11

2.2 Intrinsic parameter redundancies and estimable parameter combinations for mark-recovery y/z models 48

2.3a Intrinsic parameter redundancies for mark-recovery $x^1/y/z$ models where animals are considered as juveniles for a single year 55

2.3b Intrinsic parameter redundancies for mark-recovery $\mathrm{C}^J/y/z$ models for a general number of juvenile years J 56

2.3c Intrinsic parameter redundancies for mark-recovery $\mathrm{T}^J/y/z$ models for a general number of juvenile years J 57

2.4 Estimable parameter combinations for mark-recovery $x^J/y/z$ models . 58

3.1 An m-array representation of data from a capture-recapture study of female Greater Flamingoes from Lebreton et al. (1992) 64

3.2a All the possible capture-histories in capture-recapture models in binary notation 72

3.2b All the possible capture-history probabilities corresponding to Table 3.2a 73

3.3 Intrinsic parameter redundancies for capture-recapture y/z models . . . 77

3.4 Estimable parameter combinations for capture-recapture y/z models . . 78

3.5a Intrinsic parameter redundancies for capture-recapture $x^1/y/z$ models with first-year survival probabilities 81

3.5b Intrinsic parameter redundancies for capture-recapture $x^J/y/z$ models for a general number of juvenile years J 82

3.6 Estimable parameter combinations for capture-recapture $x^J/y/z$ models 83

3.7 Capture-histories from the dippers data set of Marzolin (1988) 85

3.8 Extrinsic parameter redundancies for capture-recapture y/z models for the dippers data set of Marzolin (1988) 88

3.9 Expected number of capture-histories for medium survival probabilities and low recapture probabilities 90

3.10 Parameter redundancies for capture-recapture y/z models where there is a maximum difference c between the number of occasions between first capture and last recapture 92

4.1 A summary of the capture-recapture-recovery model probabilities for $n_1 = n_2 = 4$ from McCrea et al. (2013) 97

4.2 All the possible life-histories in the capture-recapture-recovery model for $n_1 = n_2 \leq 4$ 104

4.3a Intrinsic parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models (Table 1) 112

4.3b Intrinsic parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models (Table 2) 113

4.4 List of confounded parameter combinations for capture-recapture-recovery $y/(z_1; z_2)$ models 114

4.5a Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1; z_2)$ models with first-year survival probabilities (Table 1) 115

4.5b Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1; z_2)$ models with first-year survival probabilities (Table 2) 116

4.5c Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1; z_2)$ models with first-year survival probabilities (Table 3) 117

4.6a Extrinsic parameter redundancy results for capture-recapture-recovery $y/(z_1; z_2)$ models for the cormorants data set of Hènaux et al. (2007) (Table 1) 121

4.6b Extrinsic parameter redundancy results for capture-recapture-recovery $y/(z_1; z_2)$ models for the cormorants data set of Hènaux et al. (2007) (Table 2) 122

4.7a Parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models where there is a maximum difference c between the number of occasions between first capture and last recapture/recovery (Table 1) 126

4.7b Parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models where there is a maximum difference c between the number of occasions between first capture and last recapture/recovery (Table 2) 127

4.8 Intrinsic parameter redundancies for capture-recapture-recovery fidelity models 135

4.9 Extrinsic parameter redundancies for capture-recapture-recovery fidelity models for the mountain chicken frogs data set 141

5.1 Extrinsic parameter redundancies in multiple-seasons occupancy models exploring the house finches data set from MacKenzie et al. (2006) . . . 159

5.2 Extrinsic parameter redundancies in multiple-groups occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012) where all histories with missing observations are ignored 165

5.3 Extrinsic parameter redundancies in multiple-groups occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012) where all histories with missing observations are now included . 171

5.4 The probabilities of occupancy-histories being observed for a certain number of detections during the history, with different ranges of low, medium and high occupancy and detection probabilities 174

5.5 Parameter redundancies for the occupancy-histories in a 2-season 3-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities 175

5.6a Parameter redundancies for the occupancy-histories in a 3-season 2-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities (Table 1) 176

5.6b Parameter redundancies for the occupancy-histories in a 3-season 2-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities (Table 2) 177

6.1 The different occupancy models examined for parameter redundancy in this thesis 180

6.2 The probability of a species true state given its observed state 188

6.3 Parameter redundancies for multiple-states occupancy models exploring the green frogs data set of MacKenzie et al. (2009) 199

6.4 Parameter redundancies in two-species interaction occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012) 215

6.5 Parameter redundancies for Royle-Nichols models for the woodthrush and catbird data sets from Fiske et al. (2014) 224

List of Supplementary Maple Files

These `Maple` files are used to display `Maple` procedures from this thesis and show parameter redundancy results in more depth. The files were created using `Maple` version 18 and more information on the software package can be found online at `http://www.maplesoft.com/products/maple`. These `Maple` files can be found attached to this thesis by USB where appropriate, or alternatively on the website `http://www.kent.ac.uk/smsas/personal/djc24/benhubbardthesis` where all the files relating to this thesis will be kept for future reference.

- `example2.1.mw`: **Example 2.1 on the mark-recovery T/A model**: This code is from Section 2.2 and begins on page 17, where it is also revisited during further examples on pages 19, 24, 32, 42 and 44.
- `example2.2.mw`: **Example 2.2 on the mark-recovery T/T model**: This code is from Section 2.2 and begins on page 20, where it is also revisited on page 37.
- `example2.3.mw`: **Example 2.3 on the mark-recovery A/T model**: This code is from Section 2.3 and begins on page 27, where it is also revisited during further examples on pages 39 and 45.
- `example2.5.mw`: **Example 2.5 on the mark-recovery $\mathbf{T}^1$/C/$\mathbf{A}_{1:2}$,T model**: This code is from Section 2.10 and begins on page 51.
- `example3.2.mw`: **Example 3.2 on the m-array capture-recapture T/T model**: This code is from Section 3.2 and begins on page 66.
- `example3.6.mw`: **Example 3.6 on the capture-recapture A,T/A,T intrinsic model**: This code is from Section 3.4 and begins on page 74.

- `example3.7.mw`: **Example 3.7 on the capture-recapture $\mathbf{T^1/C/A,T}$ intrinsic model**: This code is from Section 3.6 and begins on page 78.

- `example3.8.mw`: **Example 3.8 on the extrinsic parameter redundancy of capture-recapture models using the dippers data set**: This code is from Section 3.7 and begins on page 84, and is revisited on page 88.

- `example4.4.mw`: **Example 4.4 on the capture-recapture-recovery A/(A,T;T) intrinsic model**: This code is from Section 4.4 and begins on page 105.

- `example4.5.mw`: **Example 4.5 on the capture-recapture-recovery $\mathbf{T^1/A/(A,T;C)}$ intrinsic model**: This code is from Section 4.4 and begins on page 107.

- `example4.6.mw`: **Example 4.6 on the extrinsic parameter redundancy of capture-recapture-recovery models**: This code is from Section 4.6 and begins on page 118, and is revisited on page 123. This code is then used to analyse a data set on cormorants in Example 4.7 on page 119.

- `example4.14.mw`: **Example 4.14 on the extrinsic parameter redundancy of capture-recapture-recovery models involving fidelity**: This code is from Section 4.10 and begins on page 136. This code is then used to analyse a data set on mountain chicken frogs in Example 4.15 on page 139.

- `example5.5.mw`: **Example 5.5 on the multiple seasons occupancy model**: This code is from Section 5.6 and begins on page 156. This code is then used to analyse a data set on house finches in Example 5.6 on page 157.

- `example5.7.mw`: **Example 5.7 on the multiple seasons occupancy model with group-dependency**: This code is from Section 5.7 and begins on page 161. This code is then used to analyse a data set on amphibian breeding in Yellowstone and Grand Teton national parks in Example 5.8 on pages 163 and 170.

- `example6.8.mw`: **Example 6.8 on the multiple states occupancy model**: This code is from Section 6.1.4 and begins on page 193, and it goes further to examine a data set on green frogs in Maryland in Example 6.10 on page 197.

- `example6.9.mw`: **Example 6.9 on the multiple states occupancy model with group-dependency**: This code is from Section 6.1.4 and begins on page 195.

- `example6.11.mw`: **Example 6.11 on the initial state occupancy model**: This code is from Section 6.2 and begins on page 200.

- `example6.16.mw`: **Example 6.16 on the two-species interaction occupancy model**: This code is from Section 6.3.3 and begins on page 211, and it goes further to examine a data set on amphibian breeding in Yellowstone and Grand Teton national parks in Example 6.17 on page 214.

- `example6.20.mw`: **Example 6.20 on the Royle-Nichols occupancy model including the catbird and woodthrush data sets**: This code is from Section 6.4.3 and begins on page 220.

Chapter 1

Introduction

Statistical ecology is a widely developing field and the amount of research being undertaken on this subject is steadily increasing. This is because there is an increasing demand to study data collected on wild animals in order to make predictions about a species' future. The state of current statistical ecology research is reviewed in King (2014). We might for example be interested in animal migration rates, or how animals change location over time, or how population abundance might fluctuate in the future. These issues are important to many observers with an ever-growing need to protect, restore and manage wildlife species across the world. It is paramount that we monitor the underlying species' demographic parameters so that we can apply the correct management programs and if necessary animals can be conserved. An example of this in Great Britain is the decline in cod stocks in the Atlantic Ocean over the last couple of decades, which continues to be reported in the national press; how the government is monitoring these cod stocks is discussed in Birt et al. (2009).

Appropriate inferences are based on fitting probability models, of which there is a wide range to choose from. A potential problem when model fitting is parameter redundancy, which occurs when parameters are confounded in the model so that the model could be reparameterised in terms of a smaller number of parameters. Two examples of this include instances where the last recapture and survival parameters are confounded in the fully time-dependent capture-recapture Cormack-Jolly-Seber model (Cormack, 1964, Jolly, 1965, and Seber, 1965), and where there is age-dependent survival probability and a constant recovery probability in the Seber mark-recovery model (Seber, 1971). This is a problem we need to avoid as it makes inference impossible for some model parameters, and obtaining accurate estimates of parameters of key ecological importance is the main reason for using these models. This can be overcome

by the use of parameter constraints or by combining with other models or data sets; this however is not always obvious or possible to do in practice.

Parameter redundancy can be caused, or be made worse, due to sparse data which may arise from small sample sizes with small probabilities of detection. It is also helpful from a design perspective to be able to ascertain how rich the data set needs to be to avoid parameter redundancy. The topics considered in this thesis will help experimenters in two ways; firstly to consider which models for analysis where every parameter is in theory estimable in the model after they collected their data; and secondly to give approximate sample sizes so that full inference from the model can be made without sparse data set being an issue.

As a final example of the motivation for this thesis, we quote Professor Kenneth H. Pollock, who was the honorary speaker at the EURING Analytical Meeting 2013 in Athens, Georgia, USA. He described, "This more theoretical work [on parameter redundancy], I think, [is] incredibly important".

1.1 Introduction to parameter redundancy

This section provides a background on detecting parameter redundancy but only touches upon the actual theory and method; Chapter 2 illustrates how parameter redundancy is determined in practice and gives formal definitions.

We begin our introduction by considering **parameter identifiability**. A formal definition of parameter identifiability can be found in Silvey (1975) as well as in Catchpole and Morgan (1997) where they define a model's parameters to be identifiable if, "no two values of the parameters give the same probability distribution for the data". A model is non-identifiable if different sets of parameter values result in the same value of the likelihood. In essence, non-identifiability can occur when a model is written as a function of too many parameters, i.e. when the model is overparameterised. Silvey (1975) mentions that the problems with identifiability in model set-up is more of an *'irritant'* rather than a particular problem, though we believe however that identifiability is far more than an irritant as the problems of identifiability could mean statistical inference may be unobtainable or incorrect for certain analyses. We expand on how inference for some parameters can be unobtainable in Chapter 2.

One of the early references in the analysis of identifiability is Koopmans et al.

(1950) in which they examine whether a system of linear simultaneous equations, under some parameter constraints, is identifiable. The majority of the early work in this area take their applications from econometrics, such as in Reiersøl (1950) and Fisher (1959, 1961, 1963). Rothenberg (1971) advances methods by focusing on the information matrix to identify model identifiability rather than examining special features of particular models, which is what the earlier research does. Rothenberg (1971) is also the first to consider a general method for detecting identifiability rather than considering individual models. The methods in Rothenberg (1971) form the basis of the methodology explored in this thesis.

One area of research in which identifiability analysis has been viewed in depth is in compartment modelling. This includes compartment modelling of biological systems in Bellman and Åmström (1970) and Godfrey et al. (1982), in microbial batch growth processes in Evans and Chappell (2000), as well as in general non-linear compartment models in Walter and Pronzato (1996) and Chappell and Gunn (1998). Many of these uses of model identifiability analysis, as well as how a model's identifiability can be found in practice, can be seen in Bekker et al. (1994). Chis et al. (2011) further compares the methods of investigating parameter identifiability in the analysis of compartmental modelling. Cole et al. (2010) also shows how compartmental models can be considered under a more general framework, which includes any model that has an explicit expression. An illustrative example demonstrates how the method works for a complex linear compartmental model as seen in Audoly et al. (1998).

While there is a wide range of literature on the identifiability of compartment models, there are also other areas where identifiability has been explored. These include latent models in Goodman (1974) and van Wieringen (2005), moment structural models in Shapiro (1986), electrochemistry in Berthier et al. (1996), sludge respiration biokinetic models in Dochain et al. (1995), and naive Bayesian networks in Whiley and Titterington (2002). It is clear to see from the breadth of applications that identifiability analysis is important in many areas of statistics.

Identifiability of model parameters can be split into global and local identifiability. A model which is **globally identifiable** is one which is identifiable for every value in the parameter space, where a model which is **locally identifiable** is only identifiable for some of the areas of the parameter space. By definition, a globally identifiable model is also a locally identifiable model, but a locally identifiable model is not necessarily a globally identifiable model. In ecological models, local identifiability can occur

when there are two alternative solutions for the parameter values. An example occurs in Royle and Link (2006) when they specify the probability of an incorrect detection to be greater than a half so that the complement of that probability does not give an alternative solution. In this thesis, we mainly check whether a model's parameters are locally identifiable, which means they are only identifiable for some areas of the parameter space. Ecological models tend to be non-identifiable on boundary values such as when probabilities are equal to zero or one. Non-identifiability can also occur where there are flat ridges in likelihood surface, such as demonstrated in mark-recovery models in Catchpole and Morgan (1994, Section 3) and Catchpole and Morgan (1997, Example 3). It is further possible to make a locally identifiable model into a globally identifiable model by redefining the parameter space using parameter constraints, such as by constraining probabilities to be greater than zero and less than one.

The starting points for work on **parameter redundancy** is in Catchpole et al. (1996) and Catchpole and Morgan (1997). Parameter redundancy arises when the likelihood has flat ridges so that there is not a unique maximum likelihood parameter estimate, as proved in Catchpole and Morgan (1997, Theorem 2). This relates to the previous theory on identifiability and it can be shown that a model which is locally non-identifiable is also parameter redundant, as stated in Catchpole and Morgan (1997, Theorem 4). Catchpole and Morgan (1997) provides a symbolic method for detecting parameter redundancy in exponential family models, building on the work of Catchpole et al. (1996). Catchpole and Morgan (1997, Theorem 1) shows how to identify if a model is parameter redundant or not and further advances in Catchpole et al. (1998) demonstrate how we can identify which parameters and parameter combinations can be explicitly estimated in a parameter redundant model.

The research in this field has been significantly helped by the increase of computational power. A symbolic approach has commonly been used to determine whether a model is parameter redundant by obtaining the symbolic rank of its derivative matrix, arising from differentiating the model's terms in its specification with respect to its parameters. A symbolic algebra software package such as `Maple` can be used to perform the symbolic algebra to obtain a model's parameter redundancy. Catchpole et al. (2002) developed `Maple` procedures for detecting parameter redundancy, and these are the basis for the procedures we use in this thesis. Bailey et al. (2010) state that the symbolic approach is the, "gold standard," among methods for obtaining parameter redundancy results.

However, while the symbolic approach is the preferred one for obtaining parameter redundancy, this approach may not be computationally feasible to obtain results in `Maple` for more complex models. A numerical approach was explored in Viallefont et al. (1998) but these were found to generally be unreliable, as discussed in Gimenez et al. (2004). A joint hybrid symbolic-numerical method was developed in Choquet and Cole (2012) which combines both symbolic and numerical approaches. The development of this hybrid symbolic-numerical approach has allowed results to be found for more complex models than was possible before. A more detailed appraisal of different symbolic and numerical methods will be shown in Section 2.5.

Cole et al. (2010) extends this earlier work on parameter redundancy and gives a variety of examples of how parameter redundancy can be obtained in practice. Cole et al. (2010) provides a general framework and extends the methodology from Catchpole and Morgan (1997) for models which are not exponential family models. It also deals with finding parameter redundancy results in more complex models where other methods have previously failed. This is shown in Cole and Morgan (2010a) and Cole (2012) where Jiang et al. (2007) and Hunter and Caswell (2009) had failed to obtain results respectively. We display Tables 1.1a and 1.1b which list all of the models in ecology where parameter redundancy has been investigated. As the tables show, parameter redundancy is a hot topic of current research in statistical ecology and papers are continually being published on it. These recent papers include a wider range of more technical methods to deal with parameter redundancy such as in Cole et al. (2010) and Choquet and Cole (2012), as well as giving different parameter redundancy results for specific families of models such as in Cole et al. (2012) and Hubbard et al. (2014). It is for this reason that we write this thesis, to build on the past work in this developing field in statistical ecology.

1.2 Thesis structure

This thesis consists of five core chapters with an introductory chapter and a concluding chapter. Each core chapter relates to a different set of ecological models summarised in Tables 1.1a and 1.1b.

Chapter 2 examines parameter redundancy in mark-recovery models, where our work on parameter redundancy is based on the models shown in Seber (1970, 1971). We use mark-recovery models to demonstrate parameter redundancy theory and meth-

Table 1.1a: The different ecological models that have been examined for parameter redundancy, up to December 2014 (Table 1)

Model	Summary of model	References
Mark-recovery models	Provides survival estimates of animals where they are recovered dead	Chapter 2, Catchpole and Morgan (1996), Catchpole and Morgan (1997), Catchpole et al. (1998), Catchpole and Morgan (2001), and Cole et al. (2012)
Tag return models for fish mortality	Age-dependent models which provide estimates for mortality rates of fish	Jiang et al. (2007), and Cole and Morgan (2010a)
Ring-recovery models with age uncertainty	A ring-recovery model examining juvenile and adult animals where the number of juveniles is uncertain	Cole and Freeman (2012)
Age-dependent mark-recovery mixture models	Mixture models with mark-recovery data on animals with unknown ages when marked	McCrea et al. (2013)
Capture-recapture models	Generates survival estimates of animals from recapture occasions	Chapter 3, Catchpole and Morgan (1997), Catchpole and Morgan (2001), Catchpole et al. (2002), and Hubbard et al. (2014)
Capture-recapture mixture models	Capture-recapture models which have different classes of sub-populations	Pledger et al. (2003), and Yu et al. (2014)
Jolly-Seber and stop-over models	A capture-recapture model which allows for uncertainty in site arrival time	Matechou (2010), and Matechou and Cole: Work in progress
Multi-state capture-recapture models	Capture-recapture models where animals are in multiple states, e.g. breeding and non-breeding animals	Gimenez et al. (2003), Hunter and Caswell (2009), and Cole (2012)

ods that will be used throughout this thesis. This includes how to find if a model is parameter redundant or not, and how to obtain what parameters we can estimate if the model is parameter redundant. We give further theory including how we generalise results which have more years of marking or recovery, and how we can use different symbolic and numerical methods to obtain parameter redundancy. We use these methods to examine mark-recovery models and identify models which are not parameter

Table 1.1b: The different ecological models that have been examined for parameter redundancy, up to December 2014 (Table 2)

Model	Summary of model	References
Multi-site capture-recapture and memory models	Capture-recapture models where animals return to sites with memory of previous sites they visited	Cole et al. (2014)
Capture-recapture -recovery models	Generates survival estimates from records of alive recapture and dead recovery of animals	Chapter 4, and Hubbard et al. (2014)
Covariate models	Models where covariates are used e.g. weather survival dependency	Cole and Morgan (2010b)
Occupancy models	Generates occupancy estimates from multiple survey detections	Chapter 5, and Hubbard et al. (2015)
Multiple-states occupancy models	An occupancy model with more than a single presence state	Section 6.1
The two-species interaction model	An occupancy model with two interacting species	Section 6.3
The Royle-Nichols occupancy model	An occupancy model used to obtain estimates of animal abundance	Section 6.4
Integrated population modelling	Models for combining different independent data sets leading to one overall likelihood	Cole and McCrea (2012)
Models with individual random effects	Modelling individual random effects with the purpose of making some parameters estimable	Cole and Choquet (2012)

redundant. There is also an extension which considers mark-recovery models where there is a different set of survival parameters when the animal is a juvenile. The results listed in Chapter 2 have been published in Cole et al. (2012).

Chapter 3 considers capture-recapture models where we examine the parameter redundancy in models originating from Cormack (1964), Jolly (1965) and Seber (1965). We consider a simplified model to begin with and then investigate models using the probabilities of individual animal capture possibilities as a base for parameter redundancy analysis. We explore a variety of models where survival and recapture probabilities can be either age- and/or time-dependent. We investigate parameter redundancy

by providing a simpler set of probability combinations, and show we can also consider a capture-recapture model with a different set of juvenile survival parameters in the same manner. We then illustrate parameter redundancy where not every possible capture-history is observed by using an example data set of European dippers from Marzolin (1988). Some of the work in Chapter 3 has been published in Hubbard et al. (2014).

Chapter 4 expands the previous work on mark-recovery and capture-recapture models by considering both dead recoveries and live recaptures in the same capture-recapture-recovery model, where our work is based on the models shown in Burnham (1993), Catchpole et al. (1998) and King and Brooks (2003). Capture-recapture-recovery models considering both age- and/or time-dependent parameters are considered, as well as examining models with a different set of juvenile survival parameters. Parameter redundancy where not all of the recapture/recovery possibilities are observed is examined by exploring a data set on Great Cormorants from Hènaux et al. (2007). We also consider an extension of this model which considers the probability of fidelity, i.e. the probability of not emigrating, as an additional parameter, where our work is based on the model demonstrated in Burnham (1993). We illustrate some parameter redundancy results in this fidelity model using a data set of mountain chicken frogs. Some of this work in Chapter 4 has been published in Hubbard et al. (2014).

Chapter 5 considers occupancy models which concentrate on achieving parameter estimates for the probability of species presence or absence at particular locations. These models can be used where marking individuals, such as in the models of Chapters 2, 3 and 4, is more problematic. The basic occupancy model with only one detection survey can be extended by increasing the number of surveys and seasons, and by further considering multiple location sites, where we examine the parameter redundancy in models originating from MacKenzie et al. (2002, 2003). We give parameter redundancy results for the case where all of the possible occupancy-histories are observed. The case where not all of the possible occupancy-histories are observed is then considered in examples on house finches from MacKenzie et al. (2006) and amphibian breeding from Gould et al. (2012). We then conclude with an analysis on how approximately rich the data set needs to be so that the occupancy models used are not parameter redundant. Some of this work in Chapter 5 will be published in Hubbard et al. (2015) which is work in progress.

The final chapter, Chapter 6, considers parameter redundancy for various model extensions of the occupancy models shown in Chapter 5. This includes the addition of multiple states (such as in MacKenzie et al., 2009), the extension to a multiple number of species (such as in MacKenzie et al., 2004), and estimation of animal abundance in the Royle-Nichols occupancy model (such as in Royle and Nichols, 2003). Parameter redundancy results are illustrated for a data set of green frogs in Maryland from MacKenzie et al. (2009), a data set on amphibian breeding from Gould et al. (2012), and data sets on woodthrush and catbirds from Fiske et al. (2014) respectively.

The symbolic algebra used to investigate parameter redundancy can be executed in a symbolic algebra package such as `Maple` or `Mathematica`. We use `Maple 18` in the thesis as it follows on from earlier work on parameter redundancy in `Maple` from Catchpole and Morgan (1997), Catchpole et al. (2002) and Cole et al. (2010). We provide an electronic supplementary appendix of `Maple` files to supplement the work in this thesis. The list of the `Maple` files given as supplementary material can be found before this introductory chapter. For some of the examples given in the thesis, the words "[See electronic appendix `exampleX.Y.mw`]" indicates that there is an electronic appendix `Maple` file to supplement that example. These files give further insight into the examples and provide the `Maple` code to show how the parameter redundancy results were obtained.

Chapter 2

Mark-Recovery Models with Parameter Redundancy Examples

2.1 Mark-recovery models

The first type of models we consider for parameter redundancy analysis are **mark-recovery models**, sometimes called ring-recovery or band-recovery models depending on the study. A mark-recovery study would typically begin with the marking of animals by placing unique identifiers on them, for example attaching a small individually numbered metal or plastic tag to the leg of each bird. These identifiers are then recovered from dead animals and the data are collated on some time scale, typically yearly. A mark-recovery model can then be used to estimate survival probabilities.

Mark-recovery models originate from Haldane (1955) and Chapman and Robson (1960). Seber (1970, 1971) then formalised the previous work to create a model that used parameters to measure the recovery reporting rates and the survival probabilities during particular time periods. This is the basic model we consider in this chapter. Examples of the use of mark-recovery models can be found in Brownie et al. (1985), Krementz et al. (1997) and Seber (2002). This methodology was especially popular in the early 1980's, as demonstrated in Anderson and Burnham (1980), Nichols et al. (1982), Pollock and Raveling (1982), White (1983), Conroy and Williams (1984) and Anderson et al. (1985), to name a few. Lakhani and Newton (1983) consider a model where the survival probabilities vary with age, rather than time as explored in the papers above and we consider age-dependence parameters in our models later on in this

Table 2.1: A d-array representation of data from a mark-recovery study of lapwings from Catchpole et al. (1999)

Year of	Number of	Year of recovery								
ringing	birds ringed	1964	1965	1966	1967	1968	1969	1970	1971	1972
1963	1147	13	4	1	2	1	0	0	1	0
1964	1285		16	4	3	0	1	1	0	0
1965	1106			11	1	1	1	0	2	1
1966	1615				10	4	2	1	1	1
1967	1618					11	1	5	0	0
1968	2120						9	5	4	0
1969	2003							11	9	4
1970	1963								8	4
1971	2463									4

chapter. This work was further extended to incorporate different first year post-release survival probabilities, which is a common occurrence in many bird species, for example in Morgan and Freeman (1989), Freeman and Morgan (1990, 1992) and Catchpole et al. (1999). Unknown ages can also be incorporated in the model as a development of mixture modelling, such as in McCrea et al. (2013), but here we only consider age-dependence for animals marked as young. The main paper that this chapter will regularly refer to is Cole et al. (2012), which has all the mark-recovery parameter redundancy results given in this chapter.

The general form of data for a mark-recovery study can be represented as a **d-array** of all the animal recovery possibilities. A d-array presents the numbers of animals which have been recovered during certain recovery occasions over the course of the study. Each row represents different occasions of marking and each column represents different recovery occasions. Data in mark-recovery studies are usually displayed as a d-array for example in Catchpole et al. (1996, 1999) and Catchpole and Morgan (1997). We give an example of a d-array in Table 2.1 for a study of lapwings (*Vanellus vanellus*) ringed as nestlings, given in Catchpole et al. (1999) and Cole et al. (2012).

Let $\phi_{i,j}$ represent the probability that an animal of age $i-1$ at time j survives until time $j+1$, given that the animal has survived for all its years of life before year j since the animal was marked. For example, the parameter $\phi_{2,3}$ would indicate an animal's survival probability in its second year of life during the third year of the study. We consider age-dependency in the models of this thesis when animals are known to be of

age 0 when they marked. Further let $\lambda_{i,j}$ represent the probability that an animal of age $i-1$ is recovered dead in the jth year of the study. We refer to survival and recovery probabilities on a yearly scale in this chapter, however this does not necessarily have to be the case.

Let $N_{i,j}$ denote the number of animals marked in year i and recovered in year j for $i = 1, \ldots, n_1$ and $j = 1, \ldots, n_2$, where n_1 is the number of years of marking, and n_2 is the number of years of recovery. In this model, $n_1 \leq n_2$ as there are at least as many years of recovery as years of marking, however frequently in mark-recovery studies the number of years of marking and recovery are the same so that $n_1 = n_2$. The probability an animal was marked in year i and recovered dead in year j is

$$P_{i,j} = \left(\prod_{k=i}^{j-1} \phi_{k-i+1,k} \right) (1 - \phi_{j-i+1,j}) \cdot \lambda_{j-i+1,j}. \tag{2.1}$$

This shows that the animal has survived every year since it was marked up to year j where the animal does not survive the year and is then recovered dead. The probabilities $P_{i,j}$ can be summarised in a matrix, $\mathbf{P}$, known as its **p-array**. This p-array gives the probability of each recovery possibility corresponding to the d-array, and is given as

$$\mathbf{P} = \begin{bmatrix} (1-\phi_{1,1})\lambda_{1,1} & \phi_{1,1}(1-\phi_{2,2})\lambda_{2,2} & \phi_{1,1}\phi_{2,2}(1-\phi_{3,3})\lambda_{3,3} & \cdots \\ 0 & (1-\phi_{1,2})\lambda_{1,2} & \phi_{1,2}(1-\phi_{2,3})\lambda_{2,3} & \cdots \\ 0 & 0 & (1-\phi_{1,3})\lambda_{1,3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$\mathbf{P}$ will be of upper-triangular form as all of the terms in the entries where $i > j$ will be equal to zero as it is impossible to be recovered before being marked. This matrix is a complete summary of the mark-recovery model and it is sufficient to use just these terms to check for parameter redundancy as shown in Catchpole and Morgan (1997). Observe that if there is at least one animal recovery in every year of the study for the animals marked in year i, then we do not need to consider the possibility that not all of the animals are recovered dead during the study for the animals marked in year i as

Pr(An animal marked in year i is not recovered dead in the study)

$$= \left(1 - \sum_{j=i}^{n_2} Pr(\text{An animal marked in year } i \textit{ is} \text{ recovered dead in the study})\right)$$
$$= \left(1 - \sum_{j=i}^{n_2} P_{i,j}\right).$$

This result is due to the summation of all of the probabilities of possible animal recovery occasions being equal to one minus the probability of the animal not being recovered, as all probabilities added up must be equal to one. For this to occur, *every* year of the study must have at least one animal recovery or the result does not hold.

If we let R_i denote the number of animals marked in year i, then the likelihood for the mark-recovery model can be expressed as

$$L \propto \prod_{i=1}^{n_1} \left\{ \prod_{j=i}^{n_2} P_{i,j}^{N_{i,j}} \left(1 - \sum_{j=i}^{n_2} P_{i,j}\right)^{R_i - \sum_{j=i}^{n_2} N_{i,j}} \right\}. \tag{2.2}$$

An advantage of this mark-recovery model is that it is sufficiently flexible to allow the parameters to change in their dependencies. For example, survival and recovery probabilities that depend on age or time may be unrealistic so that these assumptions can be relaxed. We will use the same model notation that has been previously used in Catchpole and Morgan (1996) and Catchpole et al. (1996) where a model can be expressed as 'y/z', where y denotes the survival probabilities and z denotes the recovery probabilities. The probabilities of survival and recovery can then have four different alternatives:

- 'C' = All the probabilities are equal regardless of time of marking and age of animal, e.g. there is a constant survival probability ϕ throughout the study.
- 'T' = The probabilities depend on the year only, e.g. there is a different survival probability ϕ_j for each year j.
- 'A' = The probabilities depend on the age of the animal only, e.g. there is a different survival probability ϕ_i for when the animal is in its ith year of life.
- 'A,T' = The probabilities depend on the age of the animal as well as the time the animal is recovered, e.g. there are different survival probabilities $\phi_{i,j}$ dependent

on when the animal is in its ith year of life and is recovered in year j.

Various models have been used in the past for statistical inference, such as the A/C model in Seber (1971) and Catchpole and Morgan (1991), models which are age-dependent in Catchpole et al. (1996) and McCrea et al. (2013), and the models C/T, T/C and T/A in Freeman and Morgan (1992). We can generate the p-array for these mark-recovery models using the `Maple` procedure `ringmod` which is shown in the electronic supplementary appendix `Maple` files for this chapter.

We note that a slightly different parameterisation can be used specifically for tag-return models, that are more commonly used in America to describe fisheries data. This parameterisation is considered in Amstrup et al. (2005, Chapter 6), and is credited as the Brownie model (Brownie et al., 1985), in which they reparameterise the model by setting $f_{i,j} = (1 - \phi_{i,j})\lambda_{i,j}$. As we show later in Theorem 2.3, the parameterisation does not affect the model's parameter redundancy results.

The mark-recovery models of this chapter make assumptions to maintain the flexible nature and simplicity of use of these models. Violation of any of these assumptions may generate incorrect inference based on the model.

1. Marking does not affect the animal's survival probability. We do not want a situation where the marking of an animal makes it more likely to die compared to an animal which has not been marked, if we are obtaining inference about the entire animal population. An example of this assumption being violated and biased estimates coming out of the analysis is in the banding of penguins in Saraux et al. (2011), where they state that banded animals have a survival rate of 16% lower than animals which have not been banded. The consequence of this is that parameter estimates may not be truly representative of the underlying population values.

2. The discrete time period between recovery occasions needs to be same during the study. As we are working on a discrete time period this generally means yearly intervals, and these time periods have to remain at yearly intervals throughout the study.

3. Marked animals are of known age 0 when we mark them in the study, if an age-dependent model is considered. This assumption has been relaxed for animals of

unknown ages in McCrea et al. (2013) but such flexibility does increase model complexity.

4. The animals die and are recovered independently of each other.

5. Technological and/or human errors in the data are not considered, or at least are considered to be of negligible value. This includes issues such as incorrect individual identification or incorrect data input into computers.

6. The animals recovered dead in year j are assumed to have died in year j. Late reporting of recoveries could happen in practice, but this is assumed not to happen or at least be of a negligible amount.

We now identify how we can obtain the parameter redundancy of a particular model by using the mark-recovery model as defined earlier as a basis for our examples. This chapter will include how parameter redundancy can be obtained, and how the results may be generalised for larger models of the same structure such as by additional year of animal recovery. We then view various different parameter redundancy issues in Sections 2.4 to 2.8, and conclude by displaying general results for mark-recovery models in Sections 2.9 and 2.10.

2.2 What is parameter redundancy and how can it be detected?

We begin with a definition of identifiability and show how this links to parameter redundancy. Let $\mathbf{M}(\boldsymbol{\theta})$ be the function that defines a model containing the parameters of interest $\boldsymbol{\theta}$. Consider two different parameter sets $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. We then define parameter identifiability as in Definition 2.1 below.

Definition 2.1. *A statistical model,* $\mathbf{M}(\boldsymbol{\theta})$*, containing the parameters* $\boldsymbol{\theta}$*, is parameter identifiable if there is a one-to-one mapping where* $\mathbf{M}(\boldsymbol{\theta}_1) = \mathbf{M}(\boldsymbol{\theta}_2)$ *means that* $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ *for all* $\boldsymbol{\theta}_1$ *and* $\boldsymbol{\theta}_2$ *in the parameter space.*

We now give the definition of when a model is parameter redundant, as given in Definition 2.2.

Definition 2.2. *A model is parameter redundant if the parameter vector $\boldsymbol{\theta}$, of dimension t, can be replaced by a vector $\boldsymbol{\beta}$, of dimension $q < t$, such that the model can be expressed in terms of $\boldsymbol{\beta}$ alone.* (Catchpole et al., 1996, Definition 1)

The link between the two definitions is that if a model is parameter redundant, the model is *also* **parameter non-identifiable**, as in both cases different parameter values give the same model. This is shown in Catchpole and Morgan (1997) and Cole et al. (2010) as if you can write the model using fewer parameters, the model is parameter redundant and is also parameter non-identifiable. If a model is *not* parameter redundant, then all the parameters in the model can be estimated in principle. The cause of parameter redundancy and parameter non-identifiability is due to the overparameterisation of the model, as shown for the mark-recovery model given in Example 2.2. Note that *parameter* identifiability is different from *model* identifiability, as model non-identifiability is where different models result in the same value of the likelihood, rather than different parameters resulting in the same value of the likelihood. We deal solely with analysis into parameter redundancy, and therefore analysis into parameter identifiability, in this thesis.

There may also be parameter redundancy where the data set is incomplete, such as where $N_{i,j} = 0$ in mark-recovery models, but for now we work with the assumption that the data set is 'perfect' and every recovery possibility is seen at least once in a data set (see Section 2.6 for more details). A parameter redundant model is an undesirable feature as we ideally want to be able to estimate every parameter in the model. For example in a mark-recovery model, if you cannot estimate any of the survival probabilities for the animals in the study, then it would be a poor model to base inference from as one of the major objectives of using that model is the ability to estimate survival probabilities for the animals in the study.

To investigate parameter redundancy of a particular model, we need a unique representation of that model. We use the term **exhaustive summary** to define a vector of parameter combinations that uniquely defines the structure of the model. This is a term borrowed from the work on identifiability in compartment modelling from Walter and Lecourtier (1982). Definition 2.3 defines an exhaustive summary, $\boldsymbol{\kappa}$, as

Definition 2.3. *A parameter vector $\boldsymbol{\kappa}(\boldsymbol{\theta})$ is an exhaustive summary if knowledge of $\boldsymbol{\kappa}(\boldsymbol{\theta})$ uniquely determines $\mathbf{M}(\boldsymbol{\theta})$.* (Cole et al., 2010, Definition 4)

In classical statistical inference, an example for what $\mathbf{M}(\boldsymbol{\theta})$ could be is the model likelihood function, though it alternatively could be a suitable probability distribution or a series of differential equations (Cole et al., 2010, Example 2). For mark-recovery models, the vector of the non-zero terms in the model's p-array is an exhaustive summary as it uniquely defines the model. We work with other exhaustive summaries in later chapters and we typically use the natural logarithm of each non-zero term as logs are generally easier to differentiate. Discussion of other exhaustive summary forms can be found in Section 2.4.

Example 2.1 - The mark-recovery T/A model: [See electronic appendix `example2.1.mw`] This example will take the form of a mark-recovery model where survival probabilities are only time-dependent and recovery probabilities are only age-dependent. In its matrix form this model for three years of marking and recovery can be given by the `Maple` code

```
P := ringmod(2,3,3,3);
# Inputs: (y,z,n1,n2);
# y = survival probability; z = recovery probability;
# for y and z: '1'=C, '2'=T, '3'=A, '4'=A,T;
# n1 = number of years of marking; n2 = number of years of recovery.
```

The `Maple` procedure `ringmod` returns the p-array as

$$\mathbf{P} = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_2)\lambda_2 & \phi_1\phi_2(1-\phi_3)\lambda_3 \\ 0 & (1-\phi_2)\lambda_1 & \phi_2(1-\phi_3)\lambda_2 \\ 0 & 0 & (1-\phi_3)\lambda_1 \end{bmatrix}.$$

This can be used as an exhaustive summary by using only the non-zero terms in $\mathbf{P}$, as obtained by using the `Maple` procedure `Matvec`,

```
kappa := Matvec(P);
# Matvec(P) gives a vector of all non-zero exhaustive
#  summary terms from the matrix P.
```

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_2)\lambda_1 \\ \phi_2(1-\phi_3)\lambda_2 \\ (1-\phi_3)\lambda_1 \end{bmatrix}.$$

This gives us an exhaustive summary we can use to check the model for parameter redundancy. □

If a model is parameter redundant then not all the parameters can be estimated. However, while some of the original parameters themselves may not be estimable, some combinations of these parameters will still be able to be obtained. We define these as the **estimable parameter combinations** in the model and we can determine how many estimable parameter combinations we can find in a model from Theorem 2.1. We firstly define some terminology, beginning with **parameter deficiency** in Definition 2.4 below.

Definition 2.4. *The parameter deficiency of a model gives the difference between how many parameters there are in the model and the number of estimable parameter combinations that can actually be obtained.*

We give Theorem 2.1 later which shows how a model's parameter deficiency can be obtained. If the number of parameters is equal to the number of the estimable parameter combinations, then all of the original parameters can be estimated and the model is not parameter redundant. When the number of estimable parameters is equal to the number of parameters in the model, it is said to be **full rank**.

Let the parameter set of $\boldsymbol{\kappa}$ containing the p parameters in the model be denoted by $\boldsymbol{\theta}$. From the exhaustive summary $\boldsymbol{\kappa}$, we can construct the **derivative matrix** as in Definition 2.5 as

Definition 2.5. *The derivative matrix of a model, where κ_j is the jth element of the exhaustive summary $\boldsymbol{\kappa}$ for $j = 1, \ldots, r$, and θ_i is the ith parameter in the parameter set $\boldsymbol{\theta}$ for $i = 1, \ldots, p$, is defined as*

$$\mathbf{D} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} \frac{\partial \kappa_1}{\partial \theta_1} & \frac{\partial \kappa_2}{\partial \theta_1} & \cdots & \frac{\partial \kappa_r}{\partial \theta_1} \\ \frac{\partial \kappa_1}{\partial \theta_2} & \frac{\partial \kappa_2}{\partial \theta_2} & \cdots & \frac{\partial \kappa_r}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \kappa_1}{\partial \theta_p} & \frac{\partial \kappa_2}{\partial \theta_p} & \cdots & \frac{\partial \kappa_r}{\partial \theta_p} \end{bmatrix}.$$

This definition leads to Theorem 2.1 to determine if a model is parameter redundant or not.

Theorem 2.1. *a. (i) If* $\boldsymbol{D}$ *has rank equal to* p *then the model is full rank.*
a. (ii) If the rank of $\boldsymbol{D}$ *is equal to* $q < p$*, then the model is parameter redundant. There are* q *estimable parameters and the model has parameter deficiency* $d = p - q$*.*
b. If the model is parameter redundant the estimable parameters can be determined by solving $\boldsymbol{\alpha}(\boldsymbol{\theta})^T \boldsymbol{D}(\boldsymbol{\theta}) = 0$*, which has* d *solutions, labelled* $\alpha_j(\boldsymbol{\theta})$ *for* $j = 1, \ldots, d$*, with individual entries* $\alpha_{ij}(\boldsymbol{\theta})$*. Any* $\alpha_{ij}(\boldsymbol{\theta})$ *which are zero for all* d *solutions correspond to a parameter,* θ_i*, which is estimable. The solutions of the system of linear first-order partial differential equations (PDEs),*

$$\sum_{i=1}^{p} \alpha_{ij} \frac{\partial \psi}{\partial \theta_i} = 0,$$

for $j = 1, \ldots, d$ *(with* ψ *an arbitrary function), form the set of estimable parameters. Parameterised in terms of the estimable parameters, the model is full rank.* (Cole et al., 2010, Theorem 2)

Part a. of Theorem 2.1 shows how to check whether a model is parameter redundant or not. Part b. of Theorem 2.1 shows how a set of estimable parameters can be found, if the original model is parameter redundant. Catchpole et al. (1998) use the same method of solving partial differential equations in exponential family models. Chappell and Gunn (1998) and Evans and Chappell (2000) also use this same method for finding estimable parameter combinations in compartment models. We show how we can use Theorem 2.1 to identify if a model is parameter redundant or not by revisiting Example 2.1 below.

Example 2.1 revisited: [See electronic appendix `example2.1.mw`] We can first list the parameter set of the T/A mark-recovery model by the `Maple` procedure `parsproc` as

```
theta := parsproc(kappa);
# parsproc(P) gives a vector of all parameters in exhaustive summary kappa.
```

$$\boldsymbol{\theta} = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T.$$

The derivative matrix of the exhaustive summary given its parameters can then be found by the `Maple` procedure `Dmat` as

```
DD := Dmat(kappa,theta);
# Dmat(kappa,theta) generates the derivative matrix of the exhaustive
#  summary, kappa, differentiated by all of the parameters, theta.
```

$$\mathbf{DD} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} -\lambda_1 & (1-\phi_2)\lambda_2 & \phi_2(1-\phi_3)\lambda_3 & 0 & 0 & 0 \\ 0 & -\phi_1\lambda_2 & \phi_1(1-\phi_3)\lambda_3 & -\lambda_1 & (1-\phi_3)\lambda_2 & 0 \\ 0 & 0 & -\phi_1\phi_2\lambda_3 & 0 & -\phi_2\lambda_2 & -\lambda_1 \\ (1-\phi_1) & 0 & 0 & (1-\phi_2) & 0 & (1-\phi_3) \\ 0 & \phi_1(1-\phi_2) & 0 & 0 & \phi_2(1-\phi_3) & 0 \\ 0 & 0 & \phi_1\phi_2(1-\phi_3) & 0 & 0 & 0 \end{bmatrix}.$$

Note that we use **DD** to represent a model's derivative matrix as **D** is a protected variable in `Maple`. We can find the symbolic rank of this matrix using the built in `Rank` function in `Maple`, as well as finding the parameter deficiency using the built in `Dimension` function. For this derivative matrix we find that the matrix rank and model parameter deficiency are equal to

```
r := Rank(DD); d := Dimension(theta)-r;
```

$\mathbf{r} := \mathbf{6}, \mathbf{d} := \mathbf{0}$.

The output shows the model rank for the T/A mark-recovery model is equal to 6. We then see that the parameter deficiency of the model is equal to 0 as the number of parameters in the model is 6 so that $d = p - q = 0$. By part a.(i) of Theorem 2.1 the model is full rank and it is theoretically possible to estimate all of the parameters. □

We now consider another mark-recovery model in Example 2.2 and identify if that model is parameter redundant or not.

Example 2.2 - The mark-recovery T/T model: [See electronic appendix `example2.2.mw`] This is a similar example to Example 2.1 but now the recovery probabilities are time-dependent rather than age-dependent. For three years of marking and recovery we can check whether the model is parameter redundant using the same `Maple` code as Example 2.1:

```
P := ringmod(2,2,3,3);
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta);
r := Rank(DD); d := Dimension(theta)-r;
```

The output from `Maple` shows that the model's p-array is given as

$$\mathbf{P} = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_2)\lambda_2 & \phi_1\phi_2(1-\phi_3)\lambda_3 \\ 0 & (1-\phi_2)\lambda_2 & \phi_2(1-\phi_3)\lambda_3 \\ 0 & 0 & (1-\phi_3)\lambda_3 \end{bmatrix}.$$

The exhaustive summary used for parameter redundancy analysis is

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_2)\lambda_2 \\ \phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_3)\lambda_3 \end{bmatrix},$$

with model parameters $\boldsymbol{\theta} = [\phi_1,\, \phi_2,\, \phi_3,\, \lambda_1,\, \lambda_2,\, \lambda_3]^T$ and derivative matrix

$$\mathbf{DD} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} -\lambda_1 & (1-\phi_2)\lambda_2 & \phi_2(1-\phi_3)\lambda_3 & 0 & 0 & 0 \\ 0 & -\phi_1\lambda_2 & \phi_1(1-\phi_3)\lambda_3 & -\lambda_2 & (1-\phi_3)\lambda_3 & 0 \\ 0 & 0 & -\phi_1\phi_2\lambda_3 & 0 & -\phi_2\lambda_3 & -\lambda_3 \\ (1-\phi_1) & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_1(1-\phi_2) & 0 & (1-\phi_2) & 0 & 0 \\ 0 & 0 & \phi_1\phi_2(1-\phi_3) & 0 & \phi_2(1-\phi_3) & (1-\phi_3) \end{bmatrix}.$$

The rank and parameter deficiency of the model are equal to

$\mathbf{r} := \mathbf{5}, \mathbf{d} := \mathbf{1}$.

This derivative matrix only has rank 5 and is therefore *not* full rank as there are 6 parameters in the model. This then means, by part a.(ii) of Theorem 2.1, that there are only 5 estimable parameters in the model. To find a set of estimable parameter combinations, we have to use part b. of Theorem 2.1.

To find all of the estimable parameters we need to solve

$$\boldsymbol{\alpha}(\boldsymbol{\theta})^T \mathbf{D}(\boldsymbol{\theta}) = 0, \tag{2.3}$$

to obtain the d non-zero solutions denoted as $\alpha_j(\boldsymbol{\theta})^T$, with individual entries $\alpha_{ij}(\boldsymbol{\theta})$. This can be done by the `Maple` procedure below:

```
alpha := Estpars(DD,theta);
# Estpars(DD,theta) generates the alpha vector as well as the
#  estimable parameters given derivative matrix DD with parameters theta.
```

The code `Estpars` gives three separate outputs: The first output gives the $\boldsymbol{\alpha}$ vector as defined in Theorem 2.1:

$$\boldsymbol{\alpha} = \begin{bmatrix} 0 & 0 & (1-\phi_3)/\lambda_3 & 0 & 0 & 1 \end{bmatrix}.$$

If $\alpha_{ij}(\boldsymbol{\theta}) = 0$, then the corresponding parameter θ_i can be estimated. As $\boldsymbol{\theta} = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T$, this shows we can estimate the parameters $\phi_1, \phi_2, \lambda_1$ and λ_2, but we cannot estimate the parameters ϕ_3 and λ_3 individually. The second output in `Estpars` gives the system of linear first-order Lagrange partial differential equations

$$\sum_{i=1}^{p} \alpha_{ij} \frac{\partial \psi}{\partial \theta_i} = 0, \tag{2.4}$$

for the $j = 1, \ldots, d$ solutions from which the estimable parameter combinations can be obtained. As there is only one solution in this case, there is only a single partial differential equation to solve here,

$$\frac{\partial}{\partial \phi_3} \cdot \frac{(1-\phi_3)}{\lambda_3} + \frac{\partial}{\partial \lambda_3} \cdot 1 = 0.$$

The final output in the `Maple` code `Estpars` gives the estimable parameter combinations. This displays the final result that we estimate the parameters $\phi_1, \phi_2, \lambda_1$ and λ_2 explicitly as well as the estimable parameter combination $(1-\phi_3)\lambda_3$. If we consider the model where $s = (1-\phi_3)\lambda_3$, then the mark-recovery model given as

$$\mathbf{P}_s = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_2)\lambda_2 & \phi_1\phi_2 s \\ 0 & (1-\phi_2)\lambda_2 & \phi_2 s \\ 0 & 0 & s \end{bmatrix},$$

is not parameter redundant so that all the parameters ϕ_1, ϕ_2, λ_1, λ_2 and s can be estimated. Therefore the final result is that the T/T model with three years of marking and recovery has a parameter deficiency of $d = 1$ and the estimable parameters are $\phi_1, \phi_2, \lambda_1, \lambda_2$ and $(1-\phi_3)\lambda_3$. □

The symbolic method detailed in this section to find the parameter deficiency applies to any model with a defined exhaustive summary. Similar `Maple` procedures are used throughout this thesis. We now show how these parameter redundancy results

for any model can be generalised for any number of years of marking and recovery in the next section.

2.3 Methods for finding general parameter redundancy results

Theorem 2.1 shows how we can identify if a model is parameter redundant or not for a certain number of years of marking and recovery. However it is not possible to use the `Maple` procedures to find the parameter redundancy for general n_1 and n_2 values. For example, the mark-recovery T/T model has a parameter deficiency of 1 when $n_1 = n_2 = 4$, but a parameter deficiency of 2 when $n_1 = 3$ and $n_2 = 4$. We can use the methods given in Catchpole and Morgan (1997, Theorem 6) and Cole et al. (2010, Theorem 8) to obtain more general results for models with both age- and time-dependent parameters. Observe that these methods can be used in different statistical models, and not just mark-recovery models, as we shall see throughout this thesis. We propose three different methods of parameter redundancy proof below.

Method 1: Extension theorem

The standard **extension theorem** is a way of proving that a model is not parameter redundant for all models of a given structure, given the original simpler model for lower values of n_1 and n_2 is not parameter redundant. The extension theorem is given in Catchpole and Morgan (1997, Theorem 6) and Cole et al. (2010, Theorem 3) and is given below.

We consider an exhaustive summary $\boldsymbol{\kappa}_1(\boldsymbol{\theta}_1)$ with parameters $\boldsymbol{\theta}_1$ and with derivative matrix $\mathbf{D}_1(\boldsymbol{\theta}_1) = [\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_1]$. Consider an extended exhaustive summary $\boldsymbol{\kappa} = [\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2]$ where $\boldsymbol{\kappa}_2$ are the new terms when the model is extended. For example if we consider a mark-recovery model with $n_1 = n_2 = 3$, then if we extend to consider an extra year of recoveries, then $\boldsymbol{\kappa}_2$ would be equal to the new terms from the fourth recovery year which are terms not in the original exhaustive summary. This model is extended by the addition of extra parameters $\boldsymbol{\theta}_2$ to give us $\boldsymbol{\kappa}(\boldsymbol{\theta}') = [\boldsymbol{\kappa}_1(\boldsymbol{\theta}_1), \boldsymbol{\kappa}_2(\boldsymbol{\theta}')]$ with $\boldsymbol{\theta}' = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$. The derivative matrix of the extended model is now equal to

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1(\boldsymbol{\theta}_1) & \mathbf{D}_{2,1}(\boldsymbol{\theta}_1) \\ 0 & \mathbf{D}_{2,2}(\boldsymbol{\theta}_2) \end{bmatrix},$$

with $\mathbf{D}_{2,1}(\boldsymbol{\theta}_1) = [\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_1]$ and $\mathbf{D}_{2,2}(\boldsymbol{\theta}_2) = [\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$. Then the following extension theorem of Theorem 2.2 then applies:

Theorem 2.2. *If the original model is full rank (i.e.* $\mathbf{D}_1$ *is full rank) and* $\mathbf{D}_{2,2}$ *is full rank, then the extended model is full rank also.* (Cole et al., 2010, Theorem 3)

Theorem 2.2 is proved in Catchpole and Morgan (1997). Observe that if there is only one new parameter in the extended parameter set $\boldsymbol{\theta}_2$, then $\mathbf{D}_{2,2}$ will always be full rank as $\mathbf{D}_{2,2}$ would be a row vector which has matrix rank of 1. This means that as there is only one parameter in $\boldsymbol{\theta}_2$, the theorem still applies as $\mathbf{D}_1$ and $\mathbf{D}_{2,2}$ are both full rank. This can be seen if the mark-recovery T/C model is considered. If we extend the model by an additional year of recoveries, the extended part of the exhaustive summary contains only the one new parameter ϕ_i parameter in the extended parameter set $\boldsymbol{\theta}_2$ as the recovery probabilities are constant and there is only one survival probability for each separate year of recovery. The parameter redundancy result for the mark-recovery T/C model can then be proved by the extension theorem. Furthermore, if the extended part has no new parameters then the derivative matrix is equal to

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1(\boldsymbol{\theta}_1) & \mathbf{D}_{2,1}(\boldsymbol{\theta}_1) \end{bmatrix},$$

which is always full rank if $\mathbf{D}_1(\boldsymbol{\theta}_1)$ is full rank. We call these two cases as being *trivially* proved by the extension theorem as stated in Catchpole and Morgan (1997, Remark 7). We revisit Example 2.1 below where we use Theorem 2.2 to prove that the mark-recovery T/A model is always full rank for all values of $n_1 \geq 3$ and $n_2 \geq 3$.

Example 2.1 revisited: [See electronic appendix `example2.1.mw`] We have shown in Section 2.2 that the model is not parameter redundant when $n_1 = n_2 = 3$. Using the extension theorem, we can now prove the model is not parameter redundant for *every* $n_1, n_2 \geq 3$. Here the original exhaustive summary is equal to

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_2)\lambda_1 \\ \phi_2(1-\phi_3)\lambda_2 \\ (1-\phi_3)\lambda_1 \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T$. The derivative matrix $\mathbf{D}_1(\boldsymbol{\theta}_1) = [\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_1]$ was shown earlier to be full rank on page 19. As the derivative matrix is full rank, we can use the extension theorem to prove the model is full rank for any $n_1, n_2 \geq 3$. We extend this model first by adding one year of recovery to the original model. The exhaustive summary is then $\boldsymbol{\kappa}(\boldsymbol{\theta}') = [\boldsymbol{\kappa}_1(\boldsymbol{\theta}_1), \boldsymbol{\kappa}_2(\boldsymbol{\theta}')]$, with the parameter set equal to $\boldsymbol{\theta}' = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$. The extended part of the model, $\boldsymbol{\kappa}_2$, is given by

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} \phi_1\phi_2\phi_3(1-\phi_4)\lambda_4 \\ \phi_2\phi_3(1-\phi_3)\lambda_3 \\ \phi_3(1-\phi_4)\lambda_2 \end{bmatrix},$$

with new parameters $\boldsymbol{\theta}_2 = [\phi_4, \lambda_4]^T$. The extension theorem states that if $\mathbf{D}_1$ and $\mathbf{D}_{2,2} = [\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$ are full rank, then the extended model for any larger values of n_2 is also full rank. If we construct $\mathbf{D}_{2,2}$ we find that

$$\mathbf{D}_{2,2} = \left[\frac{\partial\boldsymbol{\kappa}_2}{\partial\boldsymbol{\theta}_2}\right] = \begin{bmatrix} -\phi_1\phi_2\phi_3\lambda_4 & -\phi_2\phi_3\lambda_3 & -\phi_3\lambda_2 \\ \phi_1\phi_2\phi_3(1-\phi_4) & 0 & 0 \end{bmatrix},$$

which has rank 2. This is full rank as there are 2 new parameters in $\boldsymbol{\theta}_2$, which means that if we extend the number of recoveries, the model is still full rank.

If we then extend for an extra year of marking the only term in the extended exhaustive summary is $\boldsymbol{\kappa}_2 = (1-\phi_4)\lambda_1$. As there are no new parameters in this exhaustive summary from the original unextended model when $n_2 = 4$, an extension of the number years of marking is trivially full rank.

This therefore proves that the model is not parameter redundant for all values of $n_1, n_2 \geq 3$. □

Method 2: Reparameterisation theorem

The extension theorem does not work for every model as the original model may not be full rank to begin with. Cole et al. (2010, Remark 1) states that, "If the original model is not full rank, we first need to find a reparameterisation of the model that is full rank. Then Theorem 2.2 [the extension theorem] can be applied to the reparameterised model, so that deficiency of the general model can be determined." In essence, the key of this method is to find a reparameterisation of the existing parameters so that the number of new parameters is equal to the rank of the derivative matrix of

the reparameterised model. The method is called the **reparameterisation theorem** and is given below.

Define a reparameterisation, $\mathbf{s}$, as one where the exhaustive summary parameterised in terms of $\mathbf{s}$, $\boldsymbol{\kappa}(\mathbf{s})$, contains fewer parameters than the exhaustive summary containing the original parameters $\boldsymbol{\kappa}(\boldsymbol{\theta})$. Let the derivative matrix of the reparameterised exhaustive summary be denoted as $\mathbf{D}_s = [\partial\boldsymbol{\kappa}(\mathbf{s})/\partial\mathbf{s}]$ and as usual let p be the number of original parameters in the model. Then Theorem 2.3 applies to show how we can reparameterise a model to work out its general parameter deficiency.

Theorem 2.3. *Let rank($\mathbf{D}_s$) = r_s, and p_s = dim($\boldsymbol{s}$). Then the following hold:*
a. (i) If $r_s = p_s$, $\boldsymbol{s}$ is a reduced-form exhaustive summary. The result of Theorem 2.1 can then be applied to $\boldsymbol{s}$, to examine model parametric structure.
a. (ii) If $r_s < p_s$, then $\boldsymbol{s}$ is not a reduced-form exhaustive summary. A reduced-form exhaustive summary may be found by first solving $\boldsymbol{\alpha}^T\mathbf{D}_s = 0$ and then solving the appropriate partial differential equations as in Theorem 2.1.
b. If rank($\partial\mathbf{s}/\partial\boldsymbol{\theta}$) = p_s, the number of estimable parameters is equal to r_s. If $r_s = p$, then the model in terms of $\boldsymbol{\theta}$ is full rank. If $r_s < p$ the model in terms of $\boldsymbol{\theta}$ is parameter redundant. (Cole et al., 2010, Theorem 8)

Theorem 2.3 is proved in Cole et al. (2010). The purpose of this reparameterisation theorem is that by rewriting $\boldsymbol{\kappa}(\boldsymbol{\theta})$ in terms of the reparameterisation $\mathbf{s}$, this reparameterised exhaustive summary, $\boldsymbol{\kappa}(\mathbf{s})$, is then structurally simpler than $\boldsymbol{\kappa}(\boldsymbol{\theta})$. An ideal reparameterisation is where the derivative matrix, $\mathbf{D}_s = [\partial\boldsymbol{\kappa}(\mathbf{s})/\partial\mathbf{s}]$, is full rank as $\text{rank}(\partial\mathbf{s}/\partial\boldsymbol{\theta}) = \dim(\mathbf{s})$. An ideal reparameterisation is also one which naturally extends for larger models. This means that if we extend the original model in terms of its reparameterised parameters $\mathbf{s}$, then the extended part of the model will contain additional parameters of a similar pattern from the ones given in $\mathbf{s}$: This can be more clearly seen in Example 2.3 below.

The result of rewriting the exhaustive summary in terms of $\mathbf{s}$ can mean that the derivative matrix $\mathbf{D_s}$ for the reparameterised exhaustive summary $\boldsymbol{\kappa}(\mathbf{s})$ can be computationally simpler to obtain than the original derivative matrix $\mathbf{D}$ for the exhaustive summary $\boldsymbol{\kappa}(\boldsymbol{\theta})$. Cole et al. (2010) shows how this can be done for complex models. There are instances where the use of a reparameterisation to obtain a simpler derivative matrix has meant parameter redundancy analysis can be performed in complex models

where it otherwise would be computationally infeasible to do so: This can be seen in Cole and Morgan (2010a) and Cole (2012) to overcome the complexity in obtaining the model's derivative matrices seen in the papers Jiang et al. (2007) and Hunter and Caswell (2009) respectively.

We require a reparameterisation so that the rank of the reparameterised derivative matrix $\mathbf{D}_s$ is the same as the number of parameters in the reparameterisation $\mathbf{s}$. If the reparameterised derivative matrix $\mathbf{D}_s$ is full rank, we can then use the extension theorem of Theorem 2.2 to create general results for the reparameterised exhaustive summary $\boldsymbol{\kappa}(\mathbf{s})$. As the extension theorem can only be used on full rank models, if the reparameterised derivative matrix $\mathbf{D}_s$ is not full rank, we require a further reparameterisation so that the reparameterised derivative matrix is then full rank. It is effectively a two-step procedure where a reparameterisation is found so its derivative matrix is full rank, and then the extension theorem of Theorem 2.2 is used to obtain a model's general parameter redundancy result. This will be demonstrated using the A/T mark-recovery model in Example 2.3 below.

Example 2.3 - The mark-recovery A/T model: [See electronic appendix `example2.3.mw`] Consider the A/T mark-recovery model where the survival probabilities are only age-dependent and recovery probabilities are only time-dependent. We can use the reparameterisation theorem of Theorem 2.3 to prove the model is parameter redundant with a parameter deficiency of 1 for all values of $n_1, n_2 \geq 3$. The model's p-array for three years of marking and recovery is given as

```
P := ringmod(3,2,3,3);
```

$$\mathbf{P} = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_2)\lambda_2 & \phi_1\phi_2(1-\phi_3)\lambda_3 \\ 0 & (1-\phi_1)\lambda_2 & \phi_1(1-\phi_2)\lambda_3 \\ 0 & 0 & (1-\phi_1)\lambda_3 \end{bmatrix}.$$

We can then find its exhaustive summary, parameter set and derivative matrix as detailed in Section 2.2 by the code

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta);
r := Rank(DD); d := Dimension(theta)-r;
```

This generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_1)\lambda_2 \\ \phi_1(1-\phi_2)\lambda_3 \\ (1-\phi_1)\lambda_3 \end{bmatrix},$$

and the parameters $\boldsymbol{\theta} = [\phi_1,\, \phi_2,\, \phi_3,\, \lambda_1,\, \lambda_2,\, \lambda_3]^T$ with derivative matrix given as

$$\mathbf{DD} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} -\lambda_1 & (1-\phi_2)\lambda_2 & \phi_2(1-\phi_3)\lambda_3 & -\lambda_2 & (1-\phi_2)\lambda_3 & -\lambda_3 \\ 0 & -\phi_1\lambda_2 & \phi_1(1-\phi_3)\lambda_3 & 0 & \phi_1\lambda_3 & 0 \\ 0 & 0 & -\phi_1\phi_2\lambda_3 & 0 & 0 & 0 \\ (1-\phi_1) & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_1(1-\phi_2) & 0 & (1-\phi_1) & 0 & 0 \\ 0 & 0 & \phi_1\phi_2(1-\phi_3) & 0 & \phi_1(1-\phi_2) & (1-\phi_1) \end{bmatrix}.$$

The rank and parameter deficiency of the model are

$\mathbf{r} := \mathbf{5}, \mathbf{d} := \mathbf{1}$.

As we have shown this model is parameter redundant, we need to use the reparameterisation theorem before we can use the extension theorem to determine the parameter deficiency for general values of n_1 and n_2. A possible reparameterisation of the model $\mathbf{s}$ can be

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ (1-\phi_1)\lambda_2 \\ (1-\phi_1)\lambda_3 \\ \phi_1(1-\phi_2)/(1-\phi_1) \\ \phi_2(1-\phi_3)/(1-\phi_2) \end{bmatrix}.$$

The original model can then be defined in terms of only reparameterised parameters

as

$$\boldsymbol{\kappa}(\mathbf{s}) = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_1)\lambda_2 \\ \phi_1(1-\phi_2)\lambda_3 \\ (1-\phi_1)\lambda_3 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 s_4 \\ s_3 s_4 s_5 \\ s_2 \\ s_3 s_5 \\ s_3 \end{bmatrix},$$

which has the parameter set $\mathbf{s} = [s_1, s_2, s_3, s_4, s_5]^T$. The derivative matrix, $\mathbf{D_s}$, can then be formed as

$$\mathbf{D}_s = \left[\frac{\partial \boldsymbol{\kappa}(\mathbf{s})}{\partial \mathbf{s}}\right] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_4 & 0 & 1 & 0 & 0 \\ 0 & 0 & s_4 s_5 & 0 & s_4 & 1 \\ 0 & s_2 & s_3 s_5 & 0 & s_3 & 0 \\ 0 & 0 & s_3 s_4 & 0 & 0 & 0 \end{bmatrix},$$

which is full rank 5 as there are now 5 parameters in the model. This reparameterised model therefore is not parameter redundant and full rank. To complete the proof for this model for all $n_1, n_2 \geq 3$, we then use the standard extension theorem of Theorem 2.2 on this reparameterised model. If we add a year of recoveries to the model, we have the new parameters $t_1 = (1-\phi_1)\lambda_4$ and $t_2 = \phi_3(1-\phi_4)/(1-\phi_3)$. These are natural extensions of the reparameterisation given before in $\mathbf{s}$. The extended part of the exhaustive summary when adding a year of recoveries is equal to

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} s_4 s_5 t_1 t_2 \\ s_4 s_5 t_1 \\ s_4 t_1 \end{bmatrix}.$$

Obtaining $\mathbf{D}_{2,2}$ by differentiating $\boldsymbol{\kappa}_2$ by $\boldsymbol{\theta}_2 = [t_1, t_2]^T$, we get

$$\mathbf{D}_{2,2} = \left[\frac{\partial \boldsymbol{\kappa}_2}{\partial \boldsymbol{\theta}_2}\right] = \begin{bmatrix} s_4 s_5 t_2 & s_4 s_5 & s_4 \\ s_4 s_5 t_1 & 0 & 0 \end{bmatrix},$$

which is full rank 2. Therefore, it is possible to extend by adding extra years of recoveries. Also observe that the model is trivially full rank if we extend the number years of marking, as the only non-zero term being added from $n_1 = 3$ to $n_1 = 4$ (while keeping $n_2 = 4$) is t_1 with no new parameters being added. Therefore, we have shown that the extended reparameterised model is not parameter redundant. As we have reparameterised the model using one fewer parameter than the number of parameters

in the original parameter set, the final parameter deficiency for the A/T mark-recovery model is then $d = 1$ for all values of $n_1, n_2 \geq 3$.

One point to make regarding this method is that there can be more than one reparameterisation. In this A/T mark-recovery model, we could have instead reparameterised the model by setting

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ \lambda_2/\lambda_1 \\ \lambda_3/\lambda_2 \end{bmatrix}.$$

We would then get this reparameterised exhaustive summary,

$$\boldsymbol{\kappa}(\mathbf{s}) = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_1)\lambda_2 \\ \phi_1(1-\phi_2)\lambda_3 \\ (1-\phi_1)\lambda_3 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_1 s_4 \\ s_2 s_5 \\ s_1 s_4 s_5 \end{bmatrix}.$$

This alternative reparameterisation would work in the same way as the previous reparameterisation and it would achieve exactly the same parameter deficiency result, a result stated in Cole et al. (2010). This is also shown in the `Maple` electronic supplementary file `example2.3.mw`. □

The obvious disadvantage of this approach is that the reparameterisations may not be of any use to us. Take for example the A/T mark-recovery model; it is possible to estimate the products $(1-\phi_1)\lambda_i$ for $i = 1, \ldots, n_2$ and $\phi_j(1-\phi_{j+1})/(1-\phi_j)$ for $j = 1, \ldots, n_2 - 1$. However, we can never explicitly determine what ϕ_i or λ_i is individually for any i, which is what we were hoping to make inference from using the model. You could suggest that these reparameterisations in this example are not very helpful to us and having parameter estimates for $\mathbf{s}$ serve little purpose in our analysis. This is not always the case however, and obtaining these estimable parameter combinations could be meaningful when say formulating suitable linear constraints for a model.

Method 3: Exhaustive terms

This method is for when the number of parameters exceeds the actual number of exhaustive summary terms in a model. Where this occurs, the rank of the derivative matrix equals the number of exhaustive summary terms and therefore the parameter deficiency is equal to the number of parameters minus the number of exhaustive summary terms in the model. This method is mentioned in Cole et al. (2012) as well as shown below in Example 2.4.

Example 2.4 - The mark-recovery A,T/C model: We demonstrate this method by using a mark-recovery model where the survival probabilities are age- and time-dependent and the recovery probabilities are constant. The p-array for this model for three years of marking and recovery is given as

$$\mathbf{P} = \begin{bmatrix} (1-\phi_{1,1})\lambda & \phi_{1,1}(1-\phi_{2,2})\lambda & \phi_{1,1}\phi_{2,2}(1-\phi_{3,3})\lambda \\ 0 & (1-\phi_{1,2})\lambda & \phi_{1,2}(1-\phi_{2,3})\lambda \\ 0 & 0 & (1-\phi_{1,3})\lambda \end{bmatrix},$$

with parameter set $\boldsymbol{\theta} = [\phi_{1,1}, \phi_{1,2}, \phi_{1,3}, \phi_{2,2}, \phi_{2,3}, \phi_{3,3}, \lambda]^T$. There are only 6 exhaustive summary terms in the model with 7 different parameters. The rank of the derivative matrix is equal to the number of exhaustive summary terms in the model which is 6. This means that the parameter deficiency is equal to the number of parameters minus the number of the exhaustive summary terms, which equals $d = 7 - 6 = 1$. This can be seen as being the case generally in the mark-recovery A,T/C model for any n_1 and n_2, as there are $n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$ exhaustive summary terms and $n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1 + 1$ parameters in the model, so the parameter deficiency will always be $d = 1$ for all values of $n_1, n_2 \geq 2$. The estimable parameters will then be every entry given in the model's p-array. □

2.4 Choices of exhaustive summary

An exhaustive summary was defined in Definition 2.3 as being a unique representation of the model in question. However, there is more than one possible exhaustive summary for any given model. Cole et al. (2010) discusses the choices of exhaustive summary and gives examples of derivative matrices. We demonstrate several alternative exhaustive summaries here using the mark-recovery T/A example as an illustrative example.

Example 2.1 revisited: [See electronic appendix `example2.1.mw`] Consider the T/A mark-recovery model where $n_1 = n_2 = 3$. One of the most natural exhaustive summaries of mark-recovery models is the vector of all the non-zero terms in the p-array, which here has the form

$$\boldsymbol{\kappa}_a = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ \phi_1\phi_2(1-\phi_3)\lambda_3 \\ (1-\phi_2)\lambda_1 \\ \phi_2(1-\phi_3)\lambda_2 \\ (1-\phi_3)\lambda_1 \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_a = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T$. This generates the derivative matrix

$$\mathbf{D}_a = \left[\frac{\partial \boldsymbol{\kappa}_a}{\partial \boldsymbol{\theta}_a}\right] = \begin{bmatrix} -\lambda_1 & (1-\phi_2)\lambda_2 & \phi_2(1-\phi_3)\lambda_3 & 0 & 0 & 0 \\ 0 & -\phi_1\lambda_2 & \phi_1(1-\phi_3)\lambda_3 & -\lambda_1 & (1-\phi_3)\lambda_2 & 0 \\ 0 & 0 & -\phi_1\phi_2\lambda_3 & 0 & -\phi_2\lambda_2 & -\lambda_1 \\ (1-\phi_1) & 0 & 0 & (1-\phi_2) & 0 & (1-\phi_3) \\ 0 & \phi_1(1-\phi_2) & 0 & 0 & \phi_2(1-\phi_3) & 0 \\ 0 & 0 & \phi_1\phi_2(1-\phi_3) & 0 & 0 & 0 \end{bmatrix},$$

which has full rank 6 so the model is not parameter redundant. An alternative exhaustive summary can be formed by using the logvector form of this exhaustive summary as

$$\boldsymbol{\kappa}_b = \begin{bmatrix} \ln\{(1-\phi_1)\lambda_1\} \\ \ln\{\phi_1(1-\phi_2)\lambda_2\} \\ \ln\{\phi_1\phi_2(1-\phi_3)\lambda_3\} \\ \ln\{(1-\phi_2)\lambda_1\} \\ \ln\{\phi_2(1-\phi_3)\lambda_2\} \\ \ln\{(1-\phi_3)\lambda_1\} \end{bmatrix},$$

with the same parameter set $\boldsymbol{\theta}_b = \boldsymbol{\theta}_a$. This generates the derivative matrix

$$\mathbf{D}_b = \left[\frac{\partial \boldsymbol{\kappa}_b}{\partial \boldsymbol{\theta}_b}\right] = \begin{bmatrix} -\frac{1}{(1-\phi_1)} & \frac{1}{\phi_1} & \frac{1}{\phi_1} & 0 & 0 & 0 \\ 0 & -\frac{1}{(1-\phi_2)} & \frac{1}{\phi_2} & -\frac{1}{(1-\phi_2)} & \frac{1}{\phi_2} & 0 \\ 0 & 0 & -\frac{1}{(1-\phi_3)} & 0 & -\frac{1}{(1-\phi_3)} & -\frac{1}{(1-\phi_3)} \\ \frac{1}{\lambda_1} & 0 & 0 & \frac{1}{\lambda_1} & 0 & \frac{1}{\lambda_1} \\ 0 & \frac{1}{\lambda_2} & 0 & 0 & \frac{1}{\lambda_2} & 0 \\ 0 & 0 & \frac{1}{\lambda_3} & 0 & 0 & 0 \end{bmatrix},$$

which also has full rank 6. This has computational complexity benefits as $\mathbf{D}_b$ is struc-

turally simpler than $\mathbf{D}_a$ therefore requiring less computational memory to calculate the rank due to ease of differentiating log terms. We can also reparameterise the original model say by making $s_{i,j} = \prod_{k=1}^{j-1} \phi_k(1-\phi_j)\lambda_i$ (with $\prod_{k=1}^{0} \phi_k = 1$) to give us another alternative exhaustive summary:

$$\boldsymbol{\kappa}_c = \begin{bmatrix} s_{1,1} \\ s_{2,2} \\ s_{3,3} \\ s_{1,2} \\ s_{2,3} \\ s_{1,3} \end{bmatrix},$$

with $\boldsymbol{\theta}_c = [s_{1,1}, s_{1,2}, s_{1,3}, s_{2,2}, s_{2,3}, s_{3,3}]^T$. This generates the derivative matrix

$$\mathbf{D}_c = \left[\frac{\partial \boldsymbol{\kappa}_c}{\partial \boldsymbol{\theta}_c}\right] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

which again has full rank 6. The rank of the derivative matrix becomes very easy to compute in this example, however a different set of reparameterisations $\mathbf{s}$ is needed for each different example. A fourth option is to use the product-multinomial mean of the model as detailed further in Catchpole and Morgan (1997). This exhaustive summary includes the number of animals marked in year i as F_i, but this is known and is therefore not a parameter to be estimated. The model has the form $F_i p_{i,j}$ and $F_i(1-\sum_{j=1}^{n_2} p_{i,j})$, where $p_{i,j}$ are the terms in $\mathbf{P}$ from Section 2.2 and the $F_i(1-\sum_{j=1}^{n_2} p_{i,j})$ terms relate to the probability that an animal is never recovered dead given the animal was marked

in year i. This then equates to the exhaustive summary

$$\boldsymbol{\kappa}_d = \begin{bmatrix} F_1(1-\phi_1)\lambda_1 \\ F_1\phi_1(1-\phi_2)\lambda_2 \\ F_1\phi_1\phi_2(1-\phi_3)\lambda_3 \\ F_1\{1-(1-\phi_1)\lambda_1 - \phi_1(1-\phi_2)\lambda_2 - \phi_1\phi_2(1-\phi_3)\lambda_3\} \\ F_2(1-\phi_2)\lambda_1 \\ F_2\phi_2(1-\phi_3)\lambda_2 \\ F_2\{1-(1-\phi_2)\lambda_1 - \phi_2(1-\phi_3)\lambda_2\}F_3(1-\phi_3)\lambda_1 \\ F_3\{1-(1-\phi_3)\lambda_1\} \end{bmatrix},$$

with $\boldsymbol{\theta}_d = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T$. This generates the derivative matrix

$$\mathbf{D}_d = \left[\frac{\partial \boldsymbol{\kappa}_d}{\partial \boldsymbol{\theta}_d}\right] = \begin{bmatrix} -F_1\lambda_1 & F_1(1-\phi_2)\lambda_2 & \dots & 0 & 0 \\ 0 & -F_1\phi_1\lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & -F_3\lambda_1 & F_3\lambda_1 \\ F_1(1-\phi_1) & 0 & \dots & F_3(1-\phi_3) & -F_3(1-\phi_3) \\ 0 & F_1\phi_1(1-\phi_2) & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

which again has full rank 6. Catchpole and Morgan (1997) did note that the probability of animals never recovered did not need to be included as exhaustive summary terms. A fifth option here to note that the derivative matrix can be appropriately scaled by using a matrix with the parameters on the diagonal and zeros otherwise, as suggested

in Catchpole and Morgan (2001) so that the derivative matrix is equal to

$$\begin{aligned}
\mathbf{D}_e &= \operatorname{diag}(\boldsymbol{\theta}_b) \cdot \left[\frac{\partial \boldsymbol{\kappa}_b}{\partial \boldsymbol{\theta}_b}\right] \\
&= \begin{bmatrix} \phi_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{(1-\phi_1)} & \frac{1}{\phi_1} & \dots & 0 & 0 \\ 0 & -\frac{1}{(1-\phi_2)} & \dots & \frac{1}{\phi_2} & 0 \\ 0 & 0 & \dots & -\frac{1}{(1-\phi_3)} & -\frac{1}{(1-\phi_3)} \\ \frac{1}{\lambda_1} & 0 & \dots & 0 & \frac{1}{\lambda_1} \\ 0 & \frac{1}{\lambda_2} & \dots & \frac{1}{\lambda_2} & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{-\phi_1}{(1-\phi_1)} & 1 & 1 & 0 & 0 & 0 \\ 0 & \frac{-\phi_2}{(1-\phi_2)} & 1 & \frac{-\phi_2}{(1-\phi_2)} & 1 & 0 \\ 0 & 0 & \frac{-\phi_3}{(1-\phi_3)} & 0 & \frac{-\phi_3}{(1-\phi_3)} & \frac{-\phi_3}{(1-\phi_3)} \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},
\end{aligned}$$

which again has full rank 6. This derivative matrix is computationally simpler than the other derivative matrices in this section apart from the result from reparameterisation given in $\boldsymbol{\kappa}_c$. What can be noted though however, is that all the different versions of the exhaustive summary and derivative matrix give exactly the same model rank and parameter deficiency. □

In this thesis we frequently use the $\boldsymbol{\kappa}_a$ form or $\boldsymbol{\kappa}_b$ logvector form in this thesis. The logvector form exhaustive summary tends to give simpler derivative terms and is therefore quicker computationally to find the rank of the derivative matrix.

2.5 Non-symbolic methods for obtaining parameter redundancy

The symbolic methods for obtaining parameter redundancy have been implemented successfully in `Maple` for many model examples, such as in Catchpole and Morgan (1997), Catchpole et al. (2002), Gimenez et al. (2004) and Cole et al. (2010, 2012). There are however situations where this symbolic approach fails where it is not computationally feasible to use the method for complex problems as the computer runs out of memory attempting to find the rank of the derivative matrix. As the rank of a matrix in `Maple` is found by performing Gaussian elimination on the rows, if the number of

matrix terms are large and algebraically complex, Gaussian elimination may fail due to a lack of computational memory. Examples of this include Jiang et al. (2007) and Hunter and Caswell (2009). Reparameterisation may be used to solve these computational problems, as in Cole and Morgan (2010a) and Cole (2012) which helps proceed with the computational complexity which occurs in Jiang et al. (2007) and Hunter and Caswell (2009) respectively. Reparameterisations may not however be straightforward to implement or automatic to find, so different non-symbolic methods have to be explored in these cases. For some situations, numerical methods can be easier to implement than the symbolic methods, such as shown in Viallefont et al. (1998). It does however suffer from unreliability in some situations and using numerical methods alone can lead to incorrect conclusions regarding parameter redundancy, such as shown in Cole and Morgan (2010a) for the conclusions from Jiang et al. (2007). A full appraisal of methods to obtain parameter deficiencies at that time can be seen in Gimenez et al. (2004). Gimenez et al. (2004) shows the weaknesses in numerical methods and why symbolic methods are strongly preferred if they are feasible.

More recent work has been on combining numerical and symbolic methods to make a method of obtaining parameter redundancy which incorporates positive elements of both numerical and symbolic approaches. This development of a **hybrid symbolic-numerical** approach has cumulated in Choquet and Cole (2012). Let $\boldsymbol{\theta}$ be the parameter set of the model and $\mathbf{D}(\boldsymbol{\theta})$ be the derivative matrix of an exhaustive summary with respect to the parameters $\boldsymbol{\theta}$. The hybrid symbolic-numerical approach is given in Definition 2.6 below.

Definition 2.6. *The hybrid symbolic-numerical algorithm is as follows:*

Step 1: Determine the exact computation of $\boldsymbol{D(\theta)}$ *using the symbolic method.*

Step 2: Evaluate $\boldsymbol{D(\theta)}$ *at* $\boldsymbol{\theta} = \boldsymbol{\theta}^*$*, where* $\boldsymbol{\theta}^*$ *are randomly chosen numerical values.*

Step 3: Find the rank of $\boldsymbol{D(\theta^*)}$*.*

Step 4: Solve $\alpha_{ij}\boldsymbol{(\theta^*)}^T\boldsymbol{D(\theta^*)} = 0$*, where* $\alpha_{ij}\boldsymbol{(\theta^*)} \approx 0$ *corresponds to when parameter* θ_i *can be estimated.*

(Choquet and Cole, 2012, Chapter 4)

For step 4, how close $\alpha_{ij}(\boldsymbol{\theta}^*)$ is to zero depends on the criteria as discussed in Choquet and Cole (2012) which includes the accuracy of the computer precision, but as a

rough criterion anything which is lower than 1×10^{-8} is deemed to be approximately equal to zero. This hybrid symbolic-numerical algorithm is used in Choquet and Cole (2012) to illustrate examples where the symbolic approach fails to calculate the parameter deficiency while the hybrid symbolic-numerical approach succeeds. It also shows in the examples the accuracy of the results as they match the results achieved by using the symbolic approach. The benefit of using the hybrid symbolic-numerical approach is the efficiency of the calculations which reduces computational complexity and time to perform the procedures. However, the hybrid symbolic-numerical approach occasionally can specify some starting parameter values which are on boundary values, which then may lead to the incorrect rank and parameter deficiency of the model being given. It is for that reason why it is better to use the pure symbolic approach as that gives us an exact answer for the matrix rank and a model's parameter deficiency, if the rank of the derivative matrix is computationally simple. When a derivative matrix's rank becomes computationally infeasible to calculate, we will then use this hybrid symbolic-numerical approach to obtain solutions. We overcome this issue of starting values being on boundary values by using five different starting points to confirm the method obtains the correct matrix rank. An example of how the hybrid symbolic-numerical algorithm is given in Example 2.2 below.

Example 2.2 revisited: [See electronic appendix `example2.2.mw`] We revisit the mark-recovery T/T model with 3 years of marking and recovery. We demonstrate that the same parameter redundancy results from Section 2.2 can be obtained using the hybrid symbolic-numerical approach using the `Maple` procedure `Formnum`. This code gives us the model rank and parameter deficiency as well as $\boldsymbol{\alpha}$ for five different starting values of $\boldsymbol{\theta}^*$. We use five different starting points to confirm the method obtains the correct rank and parameter deficiency as occasionally one of the random starting points can be on or near a boundary which may then produce the incorrect rank and parameter deficiency. We pick the minimum value in the first column of computer output as the model rank and the maximum value in the second column as the parameter deficiency.

```
P := ringmod(2,2,3,3):
kappa := Matvec(P):
theta := parsproc(kappa):
DD := Dmat(kappa,theta):
hybrid := Formnum(DD,theta);
# Formnum(DD,theta) with DD as the derivative matrix and theta as the
#  parameters. obtains the rank, deficiency and alpha using the
```

```
#  hybrid symbolic-numerical algorithm from Choquet and Cole, 2012.
```

$$\textbf{hybrid} := \begin{bmatrix} 5 & 1 & \left\{ \begin{array}{l} 1.360175... \text{ x } 10^{-12} \\ 4.773049... \text{ x } 10^{-11} \\ 0.610871 \\ 9.316751... \text{ x } 10^{-13} \\ 4.665910... \text{ x } 10^{-11} \\ 0.791729 \end{array} \right. \\ 5 & 1 & \left\{ \begin{array}{l} 1.611903... \text{ x } 10^{-13} \\ 2.210770... \text{ x } 10^{-12} \\ -0.729689 \\ 1.498786... \text{ x } 10^{-13} \\ 1.245950... \text{ x } 10^{-12} \\ -0.683779 \end{array} \right. \\ \vdots & \vdots & \vdots \\ 5 & 1 & \left\{ \begin{array}{l} 2.963504... \text{ x } 10^{-11} \\ -6.214586... \text{ x } 10^{-11} \\ -0.350553 \\ 9.309317... \text{ x } 10^{-11} \\ -3.353169... \text{ x } 10^{-11} \\ -0.936543 \end{array} \right. \end{bmatrix}.$$

The output **hybrid** gives the model rank in the first column, the parameter deficiency in the second column, and the values given in the vector $\boldsymbol{\alpha}$ in the third column, where there are five rows in **hybrid** as we perform the hybrid symbolic-numerical approach using five different sets of starting values. This shows again that the model rank is equal to 5 with a parameter deficiency of 1. We can further see what estimable parameters are in the model by viewing $\boldsymbol{\alpha}$ in the procedure `Formnum`. For each running of the `Formnum` code, a different set of $\boldsymbol{\alpha}$ values will be generated as the starting points each time will be different, but all of the rows show that the first, second, fourth and fifth values of $\boldsymbol{\alpha}$ are approximately equal to zero; we can therefore estimate the parameters ϕ_1, ϕ_2, λ_1 and λ_2. As the third and sixth values of $\boldsymbol{\alpha}$ are not approximately equal to zero, this shows it is not possible to estimate the third and sixth parameters in the parameter set, which are the parameters ϕ_3 and λ_3. These results coincide with the results previously found on page 2.5. Note that the `Maple` procedures presented later in this thesis omit the $\boldsymbol{\alpha}$ estimates from the third column; this adjusted procedure is given as the `Maple` procedure `Formnum2`. □

2.6 Intrinsic and extrinsic parameter redundancy

We define two different scenarios when we consider parameter redundancy in this thesis:

- **Intrinsic Parameter Redundancy**: This considers the model given we have 'perfect' data where at least one observation of each possible exhaustive summary term is observed. In intrinsic parameter redundancy we are identifying parameter redundancy issues in the model itself rather than model being parameter redundant due to a lack of data.
- **Extrinsic Parameter Redundancy**: This considers the model given we have 'imperfect' data, which means some of the possible exhaustive summary terms were not observed during the study. In particular this is considered in more complex models when it is unrealistic in practice to record every possible exhaustive summary term. We say data sets are 'sparse' if there are only a few distinct exhaustive summary terms, compared to 'rich' data sets which includes a majority of the distinct exhaustive summary terms in the model. We can illustrate this by revisiting Example 2.3 below where all of the possible exhaustive summary terms are not observed during a study.

Example 2.3 revisited: [See electronic appendix `example2.3.mw`] This is a case where you cannot immediately identify the parameter redundancy of the model with imperfect data. Consider a A/T mark-recovery model with 3 years of marking and 4 years of recovery where its p-array is given by

$$\mathbf{P} = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_2)\lambda_2 & \phi_1\phi_2(1-\phi_3)\lambda_3 & \phi_1\phi_2\phi_3(1-\phi_4)\lambda_4 \\ 0 & (1-\phi_1)\lambda_2 & \phi_1(1-\phi_2)\lambda_3 & \phi_1\phi_2(1-\phi_3)\lambda_4 \\ 0 & 0 & (1-\phi_1)\lambda_3 & \phi_1(1-\phi_2)\lambda_4 \end{bmatrix}.$$

Suppose we record these observations presented as a d-array:

$$\begin{bmatrix} 17 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 12 & 1 \end{bmatrix}.$$

This data is typical of a real-life study where the probability of dead recoveries becomes close to zero for a small number of years after marking, and we suppose that none of the animals which were marked during the second year of the study were recovered. There was also at least one animal not recovered dead for each year of marking so we

need to consider these possibilities as exhaustive summary terms as well. In practice, it is important to consider the probabilities for the animals which have never been recovered and these need to be included in the exhaustive summary.

If we have perfect data it can be shown that the parameter deficiency of this model is equal to one as shown in the `Maple` file `example2.3.mw`. If we exclude the possible exhaustive summary terms which have zero observations, the exhaustive summary for this data set now turns out to be equal to

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_1)\lambda_1 \\ \phi_1(1-\phi_2)\lambda_2 \\ (1-\phi_1)\lambda_3 \\ \phi_1(1-\phi_2)\lambda_4 \\ 1-\{(1-\phi_1)\lambda_1+\phi_1(1-\phi_2)\lambda_2+\phi_1\phi_2(1-\phi_3)\lambda_3+\phi_1\phi_2\phi_3(1-\phi_4)\lambda_4\} \\ 1-\{(1-\phi_1)\lambda_2+\phi_1(1-\phi_2)\lambda_3+\phi_1\phi_2(1-\phi_3)\lambda_4\} \\ 1-\{(1-\phi_1)\lambda_3+\phi_1(1-\phi_2)\lambda_4\} \end{bmatrix},$$

where the last three exhaustive summary terms refer to the probabilities that animals are not recovered during the study for animals marked in the first, second and third years respectively. If the methods from Section 2.2 are used to obtain the parameter redundancy for this model, the `Maple` file `example2.3.mw` gives us the result

$\mathbf{r} := \mathbf{6}, \mathbf{d} := \mathbf{2}$.

The parameter deficiency of the model has now increased from one to two. So for this particular data set, there is one fewer estimable parameter combination than if we had all the possible exhaustive summary terms observed in this study. This results in some of the inference for the model parameters being lost due to the sparseness of the data set observed. □

The issue of extrinsic parameter redundancy is regularly demonstrated in this thesis along with numerous real life data sets to show parameter deficiencies for different ecological models. In practice, this is the more important form of parameter redundancy as you rarely get perfect data sets, though examining intrinsic parameter redundancy is useful when suggesting a model to use for analysis.

2.7 The PLUR decomposition and near parameter redundancy

When a model is not parameter redundant it means that in theory we can estimate all the parameters. However, a model is not necessarily full rank for all values in the parameter space and is only locally identifiable. In statistical ecology a model may be full rank so that all the parameters can be theoretically estimated, but at boundary values, such as survival probabilities being equal to zero or one, the model is parameter redundant. In practice we observe that generally this will not be the case but these values are possible, such as in the capture-recapture Soay sheep study of Catchpole et al. (2000) where the recapture probabilities are approximately equal to one. We term such parameter values where the model is not full rank as being on boundary values and more discussion on boundary estimation in mark-recovery models can be found in Catchpole and Morgan (1994) and Catchpole et al. (1996). This is similar to having flat ridges in the likelihood space for particular parameter values, such as shown in Catchpole and Morgan (1997, Example 3) where a mark-recovery model becomes parameter redundant if all the first year survival probabilities are the same regardless of year ringed. This problem has also been seen in identifiability analysis in Shapiro and Browne (1983) where they consider a point in the parameter space where their model becomes non-identifiable.

Cole et al. (2010) shows how these boundary values can be determined by using a matrix decomposition of the derivative matrix. This is the (modified) **PLUR matrix decomposition** of the derivative matrix, which in this context proves extremely useful to find out a list of parameter conditions which makes the model not full rank. The PLUR decomposition in Cole et al. (2010) can be derived from Corless and Jeffrey (1997), which is the same representation of a matrix as in the Turing factorisation of a matrix from Turing (1948). This decomposition can be used to identify areas of the parameter space where the model will not be full rank. Cole et al. (2010) define the PLUR decomposition as follows in Theorem 2.4.

Theorem 2.4. *For a full rank model, write [the derivative matrix]* $\boldsymbol{D} = \boldsymbol{PLUR}$*, where* $\boldsymbol{P}$ *is a permutation matrix,* $\boldsymbol{L}$ *is a lower triangular matrix with ones on the diagonal,* $\boldsymbol{U}$ *is an upper triangular matrix and* $\boldsymbol{R}$ *is a matrix in reduced echelon form. The model is parameter redundant at* $\boldsymbol{\theta}$ *if and only if Det(*$\boldsymbol{U}$*) = 0 at a point* $\boldsymbol{\theta}$ *and* $\boldsymbol{R}$ *is defined at* $\boldsymbol{\theta}$*.* (Cole et al., 2010, Theorem 4)

Note that if one of $\mathbf{L}$, $\mathbf{U}$ or $\mathbf{R}$ are not defined at $\boldsymbol{\theta}$, then a further PLUR decomposition of $\mathbf{D}$ evaluated at $\boldsymbol{\theta}$ is required as shown in Cole et al. (2010, Remark 2). This is shown in Example 2.1 below where $\mathbf{L}$ and $\mathbf{U}$ are not defined when $\lambda_1 = 0$: This problem is dealt with in this case by letting $\lambda_1 = 0$ in the exhaustive summary and then by calculating a further the PLUR decomposition on that exhaustive summary.

The `Maple` code for finding the PLUR decomposition is using the built in `LUDecomposition` function within `Maple`. The procedure finds the determinant of the matrix $\mathbf{U}$. This determinant can be solved to get the solutions where $\det(\mathbf{U}) = 0$, and these solutions would cause parameter redundancy in the model for those parameters values. Example 2.1 below demonstrates the use of a PLUR decomposition to examine the boundary values for a mark-recovery T/A model.

Example 2.1 revisited: [See electronic appendix `example2.1.mw`] Consider again the mark-recovery T/A model for 3 years of marking and recovery. The exhaustive summary, derivative matrix and parameter set are as given in Section 2.2. To find the PLUR decomposition of $\mathbf{U}$, the following `Maple` code can be used:

```
(pp,ll,uu,rr) := LUDecomposition(DD,output=['P','L','U1','R'])
# LUDecomposition is an built in function with DD as the derivative matrix.
P := pp; L := ll; U := uu; R := rr;
DetU := Determinant(uu);
```

This output of this code gives the following matrices

$$\mathbf{P} = \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -\frac{(1-\phi_1)}{\lambda_1} & -\frac{(1-\phi_1)(1-\phi_2)}{\phi_1\lambda_1} & -\frac{(1-\phi_1)(1-\phi_3)}{\phi_1\phi_2\lambda_1} & 1 & 0 & 0 \\ 0 & -\frac{(1-\phi_2)}{\lambda_2} & -\frac{(1-\phi_2)(1-\phi_3)}{\phi_2\lambda_2} & \frac{\phi_1\lambda_1}{(1-2\phi_1)\lambda_2} & 1 & 0 \\ 0 & 0 & -\frac{(1-\phi_3)}{\lambda_3} & 0 & -\frac{(1-2\phi_1)\lambda_2}{(2-3\phi_1)\lambda_3} & 1 \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} -\lambda_1 & \bar{\phi}_2\lambda_2 & \phi_2\bar{\phi}_3\lambda_3 & 0 & 0 & 0 \\ 0 & -\phi_1\lambda_2 & \phi_1\bar{\phi}_3\lambda_3 & -\lambda_1 & \bar{\phi}_3\lambda_2 & 0 \\ 0 & 0 & -\phi_1\phi_2\lambda_3 & 0 & -\phi_2\lambda_2 & -\lambda_1 \\ 0 & 0 & 0 & -\frac{(1-2\phi_1)\bar{\phi}_2}{\phi_1} & -\frac{\bar{\phi}_1\phi_2\bar{\phi}_3\lambda_2}{\phi_1\lambda_1} & -\frac{(1-\phi_1-\phi_1\phi_2)\bar{\phi}_3}{\phi_1\phi_2} \\ 0 & 0 & 0 & 0 & \frac{(2-3\phi_1)\phi_2\bar{\phi}_3}{(1-2\phi_1)} & -\frac{(3\phi_1\phi_2-\phi_1-\phi_2)\bar{\phi}_3\lambda_1}{(1-2\phi_1)\phi_2\lambda_2} \\ 0 & 0 & 0 & 0 & 0 & \frac{(\phi_1-\phi_2)\bar{\phi}_3\lambda_1}{(2-3\phi_1)\phi_2\lambda_3} \end{bmatrix},$$

where $\bar{\phi}_i = 1 - \phi_i$. The determinant of $\mathbf{U}$ when simplified is equal to

$$\mathbf{DetU} = \phi_1\phi_2\lambda_1^2\lambda_2(1-\phi_2)(1-\phi_3)^2(\phi_1-\phi_2).$$

Therefore when $\phi_1\phi_2\lambda_1^2\lambda_2(1-\phi_2)(1-\phi_3)^2(\phi_1-\phi_2) = 0$, the model is parament redundant for that set of parameters values. It is then possible in general to use the `Maple` function `solve` to find the solutions where the determinant of $\mathbf{U}$ is equal to zero. Many of these trivial solutions are on boundary values when probabilities are either zero or one so we effectively exclude these solutions from our analysis. One possible solution however is that the determinant of $\mathbf{U}$ is equal to zero when $\phi_1 = \phi_2$. This shows that if the first and second year survival probabilities are the same, the model is parameter redundant. This is true if you observe the model

$$\mathbf{P}_s = \begin{bmatrix} (1-\phi_1)\lambda_1 & \phi_1(1-\phi_1)\lambda_2 & \phi_1^2(1-\phi_3)\lambda_3 \\ 0 & (1-\phi_1)\lambda_1 & \phi_1(1-\phi_3)\lambda_2 \\ 0 & 0 & (1-\phi_3)\lambda_1 \end{bmatrix},$$

where $\phi_1 = \phi_2$. If we now obtain this model's parameter deficiency, it is equal to one and the model is *not* full rank anymore. Therefore by using the PLUR decomposition, we have found a region in the interior of the parameter space where the model becomes parameter redundant, which is when $\phi_1 = \phi_2$. □

The above example shows that for certain parameter values meaningful inference cannot be achieved. A similar case to this was seen in Catchpole et al. (2001) where they consider a mark-recovery model with different first-year survival and recovery probabilities. There are efforts made to avoid considering boundary values in our results for this thesis. Note that the PLUR decomposition methodology can also be used with the extension theorem to obtain general results, as shown in Cole et al. (2010).

We can further use this PLUR decomposition to illustrate when models may become **near parameter redundant**. This is when it is possible for a full rank model to behave as though it is a parameter redundant model in practice when a parameter

is very close to one of its boundary values. Near parameter redundancy is defined as in Definition 2.7 below.

Definition 2.7. *A near parameter redundant model is one that is formally full rank, but might be classed as parameter redundant by an inevitably-imprecise numerical method, because the model is very similar to a model that is parameter redundant for a particular data set.* (Cole et al., 2010, pp.21)

If a full rank model is parameter redundant at $\boldsymbol{\theta}= \boldsymbol{\theta}$‘, the model may be near parameter redundant for parameter estimates close to $\boldsymbol{\theta}= \boldsymbol{\theta}$‘. Non-symbolic methods such as detailed in Section 2.5 can be used to show this near parameter redundancy as the solutions which arise are numerically similar to a parameter redundant model. A near parameter redundant model can give poor estimates for parameters as a consequence of small eigenvalues in the Hessian matrix at the maximum likelihood, as stated in Catchpole et al. (2001).

Example 2.1 revisited: If we continue to use the mark-recovery T/A model for 3 years of marking and recovery, we found that using the PLUR matrix decomposition, we obtained

$$\mathbf{DetU} = \phi_1\phi_2\lambda_1^2\lambda_2(1-\phi_2)(1-\phi_3)^2(\phi_1-\phi_2).$$

This means that the model is *near* parameter redundant when $\phi_1 \approx \phi_2$, as well as when some of the parameters are estimated as near the boundary values of zero or one. In practice, this means that if $\phi_1 \approx \phi_2$ statistical inference may be incorrect for some of the parameter estimates obtained from the model. □

This concept of near parameter redundancy has been studied before in Catchpole et al. (2001), Nasution et al. (2004) and Bailey et al. (2010). Catchpole et al. (2001) has a full discussion of near parameter redundancy and ways this affects inference, and also gives suggestions as to how this problem can be tackled in practice when it arises from data analysis. In these cases the numerical methods and the hybrid symbolic-numerical method of Choquet and Cole (2012) are advantageous as they identify this near parameter redundancy in the model unlike the pure symbolic methods. We will not consider near parameter redundancy in this thesis but it is something to consider when examining actual data sets when these problems of near parameter redundancy can appear.

2.8 Detecting parameter redundancy in models with covariates

One way of decreasing the parameter deficiency of a model, or even removing parameter redundancy altogether, is through the use of covariates. Covariates can be used when an additional variable affects the probabilities of survival or recovery: An example of this is where weather covariates affect the survival probabilities of the animals in a study such as in North and Morgan (1979), Tilling and Sterne (1999) and Brown (2010). Cole and Morgan (2010b) examined the effects of parameter redundancy when covariates are incorporated into the model. The reason why covariates are effective can be seen returning to the mark-recovery A/T model of Example 2.3.

Example 2.3 revisited: [See electronic appendix `example2.3.mw`] We can incorporate covariates into the mark-recovery A/T model by introducing a weather effect for the recovery probabilities. The original mark-recovery A/T model with 3 years of marking and recovery has rank 5 and a parameter deficiency of 1. Suppose that the reporting probability is now equal to

$$\lambda_j = \frac{1}{1+\exp(\alpha+\beta x_j)},$$

for parameters α and β, where x_j is a given time-dependent weather covariate for year j. This logistic form is used to ensure the probability remains bounded between zero and one. The model's parameter set has now changed from $\boldsymbol{\theta} = [\phi_1, \phi_2, \phi_3, \lambda_1, \lambda_2, \lambda_3]^T$ to being $\boldsymbol{\theta}_{cov} = [\phi_1, \phi_2, \phi_3, \alpha, \beta]^T$, i.e. we have turned a 6 parameter model into a 5 parameter model. The p-array of the model with these covariates is given as

$$\mathbf{P}_{cov} = \begin{bmatrix} \frac{(1-\phi_1)}{1+\exp(\alpha+\beta x_1)} & \frac{\phi_1(1-\phi_2)}{1+\exp(\alpha+\beta x_2)} & \frac{\phi_1\phi_2(1-\phi_3)}{1+\exp(\alpha+\beta x_3)} \\ 0 & \frac{(1-\phi_1)}{1+\exp(\alpha+\beta x_2)} & \frac{\phi_1(1-\phi_2)}{1+\exp(\alpha+\beta x_3)} \\ 0 & 0 & \frac{(1-\phi_1)}{1+\exp(\alpha+\beta x_3)} \end{bmatrix}.$$

If we now find the derivative matrix of this model with respect to the new parameter

set $\boldsymbol{\theta}_{cov}$, we obtain

$$\mathbf{D}_{cov} = \begin{bmatrix} \frac{-1}{1+\exp(\alpha+\beta x_1)} & \frac{(1-\phi_2)}{1+\exp(\alpha+\beta x_2)} & \cdots & \frac{-1}{1+\exp(\alpha+\beta x_3)} \\ 0 & \frac{-\phi_1}{1+\exp(\alpha+\beta x_2)} & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \frac{-(1-\phi_1)\exp(\alpha+\beta x_1)}{\{1+\exp(\alpha+\beta x_1)\}^2} & \frac{-\phi_1(1-\phi_2)\exp(\alpha+\beta x_2)}{\{1+\exp(\alpha+\beta x_2)\}^2} & \cdots & \frac{-(1-\phi_1)\exp(\alpha+\beta x_3)}{\{1+\exp(\alpha+\beta x_3)\}^2} \\ \frac{-(1-\phi_1)x_1\exp(\alpha+\beta x_1)}{\{1+\exp(\alpha+\beta x_1)\}^2} & \frac{-\phi_1(1-\phi_2)x_2\exp(\alpha+\beta x_2)}{\{1+\exp(\alpha+\beta x_2)\}^2} & \cdots & \frac{-(1-\phi_1)x_3\exp(\alpha+\beta x_3)}{\{1+\exp(\alpha+\beta x_3)\}^2} \end{bmatrix},$$

which has a rank of 5 and is full rank. Therefore, we can eliminate the parameter redundancy in this model if we have time-dependent covariates. □

This covariate information may not be useable for every model and adding covariates does not always result in a full rank model: An example of this is in the A/A mark-recovery model when examined using an age-dependent covariate instead of a time-dependent covariate (though there generally does not tend to be any useful age-dependent covariates in practice). In practice, it can be useful to include covariates to avoid parameter redundancy, but the reparameterised parameters now offer different ecological conclusions from the original parameters in the model, i.e. in Example 2.3 above, we have gone from estimating the recovery probabilities λ_1, λ_2 and λ_3, to estimating the parameters α and β. The addition of covariates can eliminate parameter redundancy in the model, but the covariate information should be relevant to the context of the model. The use of unnecessary covariates may provide poor estimates for the model parameters as increasing the complexity of the model means there is less information in the data per parameter than before.

Cole and Morgan (2010b) have examined the effects of covariates in ecological models and they proved that the number of estimable parameters in the model with covariates is equal to $\min(p_c, q)$, where p_c is the number of parameters in the covariate model and q is the number of estimable parameters in the model without covariates. We can see this in Example 2.3 previously as $p_c = 5$ in the covariate model and $q = 6$ in the original model, so the number of estimable parameters in the model with covariates is 5. Covariates and their effect on parameter redundancy is further explored in Cole and Morgan (2010b).

2.9 Intrinsic mark-recovery model results

In this section we present intrinsic parameter redundancy results for the mark-recovery y/z models as specified in Section 2.1. These parameter redundancy results listed in Table 2.2 have been presented in Cole et al. (2012, Table 1). The results in Table 2.2 are for the 16 possible mark-recovery y/z models where y and z can either be constant (C), time-dependent (T), age-dependent (A) or age- and time-dependent (A,T). The second column of Table 2.2 refers to the rank of the model, which is the number of estimable parameters in the model. The third column refers to the parameter deficiency of the model where it is parameter redundant when $d > 0$. There are n_1 years of release and n_2 years of recovery, where $n_2 \geq 2$ with $n_2 \geq n_1$. The T/A model only has its listed rank and parameter deficiency when $n_1 \geq 3$ where the model has a parameter deficiency of 1 when $n_1 = 2$. The fourth column refers to the method of proof used from Section 2.3 to obtain the parameter redundancy result, and the final column then gives the estimable parameter combinations which uses the method of solving partial differentiation equations as detailed in Section 2.2. The results of Table 2.2 assumes that perfect data are recorded with at least one animal being observed dead at each possible recovery occasion, i.e. all $N_{i,j}$ values are non-zero.

The letter 'E' in Table 2.2 is used to denote the number of terms in the model's p-array as it is of upper triangular form with $P_{i,j} = 0$ for all $i > j$, as it is impossible to have recoveries on animals which have not been marked in the first place. Note that as

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2},$$

it can be shown that the number of terms in an p-array is equal to

$$\begin{aligned}
E &= n_1 n_2 - \sum_{i=1}^{n_1} (i-1) \\
&= n_1 n_2 - \left(\sum_{i=1}^{n_1} i - \sum_{i=1}^{n_1} 1 \right) \\
&= n_1 n_2 - \left(\left(\frac{n_1(n_1+1)}{2} \right) - n_1 \right) \\
&= n_1 n_2 - \left(\frac{n_1^2}{2} + \frac{n_1}{2} - n_1 \right) \\
&= n_1 n_2 - \frac{1}{2} n_1^2 + \frac{1}{2} n_1.
\end{aligned}$$

Table 2.2: Intrinsic parameter redundancies and estimable parameter combinations for mark-recovery y/z models

Model	Rank	Deficiency	Method	Estimable Parameter Combinations
C/C	2	0	Extension	ϕ, λ
C/T	n_2+1	0	Extension	ϕ, $\{\lambda_i\}_{i\,=\,1,\ldots,n_2}$
C/A	n_2	1	Reparameterisation	$\left\{\phi^{i-1}(1-\phi)\lambda_i\right\}_{i\,=\,1,\ldots,n_2}$
C/A,T	E	1	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
T/C	n_2+1	0	Extension	$\{\phi_i\}_{i\,=\,1,\ldots,n_2}$, λ
T/T	n_1+n_2-1	n_2-n_1+1	Reparameterisation	$\{\phi_i\}_{i\,=\,1,\ldots,n_1-1}$, $\{\lambda_i\}_{i\,=\,1,\ldots,n_1-1}$, $\phi_{n_1}(1-\phi_{n_1})\lambda_{n_1}$; if $n_2>n_1$ also: $\{\phi_i(1-\phi_{i+1})\lambda_{i+1}\}_{i\,=\,n_1+1,\ldots,n_2-1}$
T/A*	$2n_2$	0	Extension	$\{\phi_i\}_{i\,=\,1,\ldots,n_2}$, $\{\lambda_i\}_{i\,=\,1,\ldots,n_2}$
T/A,T	E	n_2	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
A/C	n_2	1	Reparameterisation	$\left\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda\right\}_{i\,=\,1,\ldots,n_2}$
A/T	$2n_2-1$	1	Reparameterisation	$\{P_{1,j}\}_{j\,=\,1,\ldots,n_2}$, $\left\{\frac{\lambda_i}{\lambda_{i-1}}\right\}_{i\,=\,2,\ldots,n_2}$
A/A	n_2	n_2	Reparameterisation	$\left\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\right\}_{i\,=\,1,\ldots,n_2}$
A/A,T	E	n_2	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
A,T/C	E	1	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
A,T/T	E	n_2	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
A,T/A	E	n_2	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
A,T/A,T	E	E	Exhaustive terms	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$

* = only valid for $n_1 \geq 3$: The rank is $2n_2-1$ and parameter deficiency is 1 when $n_1=2$, as proved by the exhaustive terms method with estimable parameters $\{P_{i,j}\}_{\substack{i\,=\,1,\,2\\ j\,=\,i,\ldots,n_2}}$; $E = n_1n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.

We can see that some models are full rank but most of the more complex models are parameter redundant. The majority of the parameter redundancy above comes from when a certain set of parameters are age- and time-dependent due to the exhaustive terms in the model. This is because if one set of parameters is age- and time-dependent, say for the recovery probabilities, there is a different $\lambda_{i,j}$ parameter for each exhaustive summary term. Even if survival probabilities are then constant, there are more parameters than data points in the study meaning there is parameter redundancy in the model. Cole et al. (2012) analyses some extrinsic parameter results given that we only have a certain number of values on each diagonal of the model's p-array, but we will not quote any of these results in this thesis. Note that tag-return models, which are used in America to describe fisheries data (as discussed in Brownie et al., 1985), have the same parameter redundancy results as above. This is due to the reparameterisation theorem of Theorem 2.3 as there is a direct reparameterisation between tag-return models and the mark-recovery models explored in Table 2.2.

2.10 Intrinsic mark-recovery models with juvenile survival probabilities

It is possible to explore a variation of this model which includes a set of juvenile survival probabilities. This is often the case in some bird species where they may have a lower survival rate in the first year of life compared to when the animal is of adult age. Some examples where this has been considered for mark-recovery models are in Morgan and Freeman (1989), Freeman and Morgan (1990, 1992) and Catchpole et al. (1999). We can also extend this idea for having different survival probabilities when the animal is a juvenile for several years rather then in just its first year of life. Mead et al. (1979) and North and Morgan (1979) provide examples in which survival probabilities of herons are age-specific for the first two or three years of life, but are constant for older birds.

We let J denote the number of years where the animal is a juvenile where the animal is an adult in year $J+1$ and beyond. We can show that the probability for a single recovery occasion where an animal was marked in year i and recovered dead in

year j is equal to

$$P_{i,j} = \begin{cases} \left(\prod_{k=i}^{j-1} \phi_{k-i+1,k}\right)(1-\phi_{j-1+1,j})\lambda_{j-i+1,j} & \text{if } j \leq J, \\ \left(\prod_{k=i}^{J} \phi_{k-i+1,k}\right)\left(\prod_{k=J+1}^{j-1} \phi_{k-i+1,k}\right)(1-\phi_{j-1+1,j})\lambda_{j-i+1,j} & \text{if } j > J, \end{cases} \tag{2.5}$$

where $\prod_{k=i}^{0} \phi_{k-i+1,k} = 1$. The first possibility corresponds to when the animal is recovered dead as a juvenile and the second possibility is when the animal is recovered dead as an adult. Note that we do not have to consider the case where the animal is not recovered dead as

$$Pr(\text{An animal is not recovered dead in the study}) = \left(1 - \sum_{j=i}^{n_2} P_{i,j}\right),$$

which allows us to exclude these terms from the exhaustive summary. We use a similar notation to the y/z mark-recovery models where the mark-recovery models with juvenile survival probabilities are denoted as $x^J/y/z$, where x denotes juvenile survival probability for J years, y denotes adult survival probability, and z denotes recovery probability with the options given previously. This is slightly different from the notation used previously in Catchpole and Morgan (1996), Catchpole et al. (1996) and Cole et al. (2012). The options for different models includes the following:

- The juvenile survival probabilities can be either constant (C) with ϕ_j for $j = 1, \ldots, J$ and then have separate adult survival probabilities, or be time-dependent (T) with $\phi_{i,j}$ for $i = 1, \ldots, n_1$ and $j = i, \ldots, J$ and then have separate adult survival probabilities.

- The adult survival probabilities have the same options as before as either being constant (C), time-dependent (T), age-dependent (A), or age- and time-dependent (A,T).

- The recovery probabilities have many different options as the parameters can be either being constant (C), time-dependent (T), age-dependent (A), age-dependent for the first J years of life with separate constant adult recovery ($A_{1:J+1}$), age- and time-dependent for the first J years of life with separate time-dependent adult recovery ($A_{1:J+1}$,T), or fully age- and time-dependent (A,T).

The examples of models used in the past for inference include the T^1/A/C model in Morgan and Freeman (1989), the T^1/A/T, C^1/A/T and C^1/C/T models in Freeman

and Morgan (1990) as well as Freeman and Morgan (1992), the $\mathrm{C}^1/\mathrm{C}/\mathrm{T}$ and $\mathrm{T}^1/\mathrm{C}/\mathrm{C}$ models in Catchpole and Morgan (1996), and $\mathrm{T}^J/\mathrm{C}/\mathrm{T}$ models in McCrea et al. (2013). The methodology to allow for different parameter dependencies is quite general here and allows for a wide scope of candidate models to be considered; for example age-dependent adult survival probabilities which have different dependencies from juvenile survival probabilities may not be considered in most realistic studies. We give an example below of how the intrinsic parameter deficiency can be found for the $\mathrm{T}^1/\mathrm{C}/\mathrm{A}$ mark-recovery model using the `Maple` procedure `ringmodJ`.

Example 2.5 - The mark-recovery $T^1/C/A_{1:2},T$ model: [See electronic appendix `example2.5.mw`] This illustrative example obtains the intrinsic parameter redundancy for a $\mathrm{T}^1/\mathrm{C}/\mathrm{A}_{1:2},\mathrm{T}$ mark-recovery model where animals are marked at age 0 and are considered to have different time-dependent first-year survival probabilities with constant adult survival probability after their first year of life, along with age- and time-dependent recovery probabilities in the first year and then time-dependent recovery probabilities from the second year onwards. The model rank and parameter deficiency for a model with four years of marking and recovery can be obtained by the following code:

```
P := ringmodJ(2,1,1,5,4,4);
# Inputs: (x,J,y,z,n1,n2);
# x = juvenile survival probability;
# J = number of years that the animal is a juvenile for;
# y = adult survival probability; z = recovery probability;
# for x and y: '1'=C, '2'=T, '3'=A, '4'=A,T (x can only be C or T);
# for z: '1'=C, '2'=T, '3'=A, '4'=A_{1:J+1}, '5'=A_{1:J+1},T, '6'=A,T;
# n1 = number of years of marking; n2 = number of years of recovery.
```

`Maple` returns the p-array as given by

$$\mathbf{P} = \begin{bmatrix} (1-\phi_{1,1})\lambda_{1,1} & \phi_{1,1}(1-\phi_a)\lambda_{a,2} & \phi_{1,1}\phi_a(1-\phi_a)\lambda_{a,3} & \phi_{1,1}\phi_a^2(1-\phi_a)\lambda_{a,4} \\ 0 & (1-\phi_{1,2})\lambda_{1,2} & \phi_{1,2}(1-\phi_a)\lambda_{a,3} & \phi_{1,2}\phi_a(1-\phi_a)\lambda_{a,4} \\ 0 & 0 & (1-\phi_{1,3})\lambda_{1,3} & \phi_{1,3}(1-\phi_a)\lambda_{a,4} \\ 0 & 0 & 0 & (1-\phi_{1,4})\lambda_{1,4} \end{bmatrix},$$

where ϕ_a denotes a constant adult survival probability and $\lambda_{a,j}$ denotes a time-dependent recovery probability where age-dependence is not considered as the animal is not a juvenile at that point. From this p-array, we can then find the parameter deficiency of the model in the way described in Section 2.2,

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta):
r := Rank(DD); d := Dimension(theta)-r;
```

The model's exhaustive summary is

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_{1,1})\lambda_{1,1} \\ \phi_{1,1}(1-\phi_a)\lambda_{a,2} \\ \phi_{1,1}\phi_a(1-\phi_a)\lambda_{a,3} \\ \phi_{1,1}\phi_a^2(1-\phi_a)\lambda_{a,4} \\ (1-\phi_{1,2})\lambda_{1,2} \\ \phi_{1,2}(1-\phi_a)\lambda_{a,3} \\ \phi_{1,2}\phi_a(1-\phi_a)\lambda_{a,4} \\ (1-\phi_{1,3})\lambda_{1,3} \\ \phi_{1,3}(1-\phi_a)\lambda_{a,4} \\ (1-\phi_{1,4})\lambda_{1,4} \end{bmatrix},$$

with parameters $\boldsymbol{\theta} = [\phi_a,\ \phi_{1,1},\ \phi_{1,2},\ \phi_{1,3},\ \phi_{1,4},\ \lambda_{1,1},\ \lambda_{1,2},\ \lambda_{1,3},\ \lambda_{1,4},\ \lambda_{a,2},\ \lambda_{a,3},\ \lambda_{a,4}]^T$. The derivative matrix is given in the `Maple` file `example2.5.mw` only to save space. This derivative matrix has the model rank and parameter deficiency

$\mathbf{r} := \mathbf{9}, \mathbf{d} := \mathbf{3}$.

Therefore, the $\mathrm{T}^1/\mathrm{C}/\mathrm{A}_{1:2}$,T mark-recovery model with four years of marking and recovery has a model rank of 9 and a parameter deficiency of 3. □

Due to the increasing complexity of the parameters involved, we first introduce a 'B' term. This term is used to calculate the number of adult survival parameters as these probabilities are when the animal is of at least age $(J+1)$. If we did not consider B in our analyses, the number of parameters would be incorrectly calculated as there could be less than E survival parameters in the model. An example of this would be the model containing constant juvenile survival parameters and age- and time-dependent adult survival parameters. The B term can be obtained by adjusting our E term so that n_1 is now equal to $(n_1 - J)$ and n_2 is now equal to $(n_2 - J)$, as

shown below:

$$\begin{aligned} E-B &= (n_1-J)(n_2-J)-\frac{1}{2}(n_1-J)^2+\frac{1}{2}(n_1-J) \\ &= n_1n_2-Jn_2-Jn_1+J^2-\frac{1}{2}n_1^2+Jn_1-\frac{1}{2}J^2+\frac{1}{2}n_1-\frac{1}{2}J \\ &= n_1n_2-\frac{1}{2}n_1^2+\frac{1}{2}n_1-Jn_2+\frac{1}{2}J^2-\frac{1}{2}J \\ &= E-Jn_2+\frac{1}{2}J^2-\frac{1}{2}J \\ &= E-\frac{1}{2}J(2n_2-J+1) \\ \implies B &= \frac{1}{2}J(2n_2-J+1). \end{aligned}$$

We further define a 'G' term which is also an adjustment for the number of survival parameters in the model, which distinguishes between two cases where the matrix either has the property $n_2 < n_1 + J$ (i.e. the model's matrix is, or is quite close to being, a square matrix) or the property where $n_2 \geq n_1 + J$. The G term when the p-array has the property $n_1 \leq n_2 + J$ is

$$G = \sum_{i=1}^{n_2-n_1-1} i = \frac{(n_2-n_1-1)^2+(n_2-n_1-1)}{2} = \frac{(n_2-n_1-1)(n_2-n_1)}{2},$$

using the summation of the first n integers formula. The G term when the p-array has the property $n_1 \geq n_2 + J$ is

$$G = \sum_{i=1}^{n_2-n_1-1} i - \sum_{i=1}^{n_2-n_1-J} i = \frac{1}{2}(J-1)(2n_2-J-2n_1).$$

Note that when $n_1 = n_2$ then $G = 0$. There are simpler results when $J = 1$ as $B = n_2$ and $G = 0$ for all values of n_1 and n_2 (as $G = \frac{1}{2}(n_2-n_1-1)^2+(n_2-n_1-1) = 0$ if $n_1 = n_2$ or $n_1 = n_2 - 1$, and $G = \frac{1}{2}(J-1)(2n_2-J-2n_1) = 0$ for $n_1 \geq n_2 + 1$ if $J = 1$). This is why we give Table 2.3a for when $J = 1$ as the results are simpler, along with the more complex cases when $J > 1$ in Tables 2.3b and 2.3c.

We show three tables for mark-recovery models with juvenile survival probabilities: Table 2.3a is for the intrinsic parameter redundancy results when $J = 1$ and Tables 2.3b and 2.3c are for the intrinsic parameter redundancy results for a general J where $1 \leq J < n_2$. In the tables, the first column denotes the model with the notation $x^J/y/z$ as described previously. The model rank and parameter deficiency are then given in columns two and three respectively. The values of n_1 and n_2 where the validity of the

parameter redundancy results are correct are given in the fourth and fifth columns of all the tables as the results do not always hold for all models when there are small values of n_1 and n_2. The method of proof is given in the fifth column of Table 2.3a and the sixth column of Tables 2.3b and 2.3c. All the results assume there are perfect data observed with at least one animal recovered dead at each recovery occasion for each year of marking. These results are given in the supplementary material of Cole et al. (2012, Table 1). Cole et al. (2012) also considers some extrinsic parameter results given that we only have a certain number of values on each diagonal of the matrix $\mathbf{P}$, but we will not quote any of these results in this thesis.

Table 2.4 further shows the estimable parameter combinations for all the $x^J/y/z$ mark-recovery models that are parameter redundant. These estimable parameter combinations form the basis of a reparameterisation theorem proof which proves the results for general n_1 and n_2 values. This is given in the supplementary material of Cole et al. (2012, Table 2).

2.11 Discussion

We have begun this thesis with a chapter giving some examples of how parameter redundancy can be obtained for ecological models, using the mark-recovery model as a base model. Every result in this thesis begins with a suitable exhaustive summary which we find the derivative matrix from, and we then calculate the rank of that matrix, either by using the symbolic method of Cole et al. (2010) or the hybrid symbolic-numerical method of Choquet and Cole (2012). If the rank is equal to the number of parameters in the model, the model is full rank and all of its parameters are in theory estimable. If the rank is less than the number of parameters then the model is parameter redundant. If a model is parameter redundant, then it can be due to the model itself (intrinsic parameter redundancy) and/or due to the sparseness of data observed (extrinsic parameter redundancy). Considering extrinsic parameter redundancy in this thesis is important as it is almost impossible to achieve perfect data for complex models in long studies. We have shown how the parameter deficiency results can be generalised using the extension theorem of Catchpole and Morgan (1997) given in Theorem 2.2, or the reparameterisation theorem of Cole et al. (2010) given in Theorem 2.3.

Table 2.3a: Intrinsic parameter redundancies for mark-recovery $x^1/y/z$ models where animals are considered as juveniles for a single year

Model	Model Rank	Deficiency	Validity $n_1 \geq$	Validity $n_2 \geq$	Method
$C^1/C/C$	3	0	2	3	1
$C^1/C/T$	n_2+2	0	2	3	1
$C^1/C/A$	n_2	2	2	2	2
$C^1/C/A_{1:2}$	3	1	2	3	2
$C^1/C/A_{1:2},T$	n_1+n_2	1	2	3	2
$C^1/C/A,T$	E	2	2	2	3
$C^1/T/C$	n_2+1	0	2	3	1
$C^1/T/T$	n_1+n_2	n_2-n_1	2	3	2
$C^1/T/A$	$2n_2-1$	1	3	4	2
$C^1/T/A_{1:2}$	n_2+1	1	2	3	2
$C^1/T/A_{1:2},T$	$\min(n_1+2n_2-3, 2n_1+n_2-2)$	$\max(2, n_2-n_1+1)$	2	2	2
$C^1/T/A,T$	E	n_2	2	2	3
$C^1/A/z$	Identical to A/z models: See Table 2.2				
$C^1/A,T/C$	$E-n_1+1$	1	2	2	3
$C^1/A,T/T$	E	n_2-n_1+1	2	2	3
$C^1/A,T/A$	$E-n_1+1$	n_2	2	2	3
$C^1/A,T/A_{1:2}$	$E-n_1+1$	2	2	2	3
$C^1/A,T/A_{1:2},T$	E	n_2	2	2	3
$C^1/A,T/A,T$	E	$E-n_1+1$	2	2	3
$T^1/C/C$	n_1+2	0	2	3	1
$T^1/C/T$	n_1+n_2+1	0	3	4	1
$T^1/C/A$	n_1+n_2	1	2	3	2
$T^1/C/A_{1:2}$	n_1+3	0	2	3	1
$T^1/C/A_{1:2},T$	$\min(2n_1+n_2-2, n_1+2n_2-3)$	$\max(2, n_1-n_2+3)$	2	2	2
$T^1/C/A,T$	E	n_1+1	2	2	3
$T^1/T/C$	n_1+n_2	0	2	2	1
$T^1/T/T$	$\min(n_1+2n_2-3, 2n_1+n_2-2)$	$\max(2, n_2-n_1+1)$	2	2	2
$T^1/T/A$	n_1+2n_2-1	0	4	5	1
$T^1/T/A_{1:2}$	n_1+n_2+1	0	3	4	1
$T^1/T/A_{1:2},T$	$\min(n_1+2n_2-3, 2n_1+n_2-2)$	$\max(n_1+1, n_2)$	2	2	2
$T^1/T/A,T$	E	n_1+n_2-1	2	2	3
$T^1/A/C$	n_1+n_2	0	2	2	1
$T^1/A/T$	n_1+2n_2-1	0	4	5	1
$T^1/A/A$	n_1+n_2	n_2-1	2	3	2
$T^1/A/A_{1:2}$	n_1+n_2	1	3	4	2
$T^1/A/A_{1:2},T$	$\min(n_1+3n_2-6, 2n_1+2n_2-5)$	$\max(n_1-n_2+4, 3)$	2	2	2
$T^1/A/A,T$	E	n_1+n_2-1	2	2	3
$T^1/A,T/z$	Identical to $A,T/z$ models: See Table 2.2				

Method 1 = Extension, Method 2 = Reparameterisation, Method 3 = Exhaustive terms;
$E = n_1n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.

Table 2.3b: Intrinsic parameter redundancies for mark-recovery $C^J/y/z$ models for a general number of juvenile years J

Model	Model Rank	Parameter Deficiency	Validity $n_1 \geq$	Validity $n_2 \geq$	Method
C^J/C/C	$J+2$	0	$J+1$	$J+2$	1
C^J/C/T	n_2+J+1	0	$J+1$	$J+2$	1
C^J/C/A	n_2	$J+1$	$J+1$	$J+1$	2
C^J/C/$A_{1:J+1}$	$J+2$	J	$J+1$	$J+2$	2
C^J/C/$A_{1:J+1}$,T	$n_1+B-J+1-G$	J	$J+1$	$J+2$	2
C^J/C/A,T	E	$J+1$	$J+1$	$J+1$	3
C^J/T/C	n_2+1	0	$J+1$	$J+2$	1
C^J/T/T	$\min(2n_2, n_1+n_2+J-1)$	$\max(0, n_2-n_1-J+1)$	$J+1$	$J+2$	1*
C^J/T/A	$2n_2-J$	J	$J+3$	$J+3$	2
C^J/T/$A_{1:J+1}$	n_2+1	J	$J+1$	$J+2$	2
C^J/T/$A_{1:J+1}$,T	$\begin{cases} B+n_1+n_2-2J-1-G & n_2 < n_1+J \\ B+2n_1-J-1-G & n_2 \geq n_1+J \end{cases}$	$\begin{cases} J+1 & n_2 < n_1+J \\ n_2-n_1+1 & n_2 \geq n_1+J \end{cases}$	$J+1$	$J+1$	2
C^J/T/A,T	E	n_2	$J+1$	$J+1$	3
C^J/A/z	Identical to A/z models: See Table 2.2				
C^J/A,T/C	$E-n_1+n_2-B+J+G$	1	$J+1$	$J+1$	2
C^J/A,T/T	$\begin{cases} E & J=1 \\ E-n_1+2n_2-B+J-1+G & J>1, n_2<n_1+J \\ E+n_2-B+2J-2+G & J>1, n_2 \geq n_1+J \end{cases}$	$\begin{cases} n_2-n_1+1 & J=1 \\ 1 & J>1, n_2<n_1+J \\ n_2-n_1-J+2 & J>1, n_2 \geq n_1+J \end{cases}$	$J+1$	$J+1$	2
C^J/A,T/A	$E-n_1+n_2-B+J+G$	n_2	$J+1$	$J+1$	2
C^J/A,T/$A_{1:J+1}$	$E-n_1+n_2-B+J+G$	$J+1$	$J+1$	$J+1$	2
C^J/A,T/$A_{1:J+1}$,T	E	n_2	$J+1$	$J+1$	3
C^J/A,T/A,T	E	$E-n_1+n_2-B+J+G$	$J+1$	$J+1$	3

Method 1 = Extension, Method 2 = Reparameterisation, Method 3 = Exhaustive terms;
*Proof is the extension theorem if $d=0$ when $n_2 > n_1+J-1$ and the reparameterisation theorem if $d>0$;
$E = n_1n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1; B = \frac{1}{2}J(2n_2-J+1)$;
$G = \frac{1}{2}[(n_2-n_1-1)^2 + (n_2-n_1-1)$ when $n_2 \leq n_1+J$, and $G = \frac{1}{2}(J-1)(2n_2-J-2n_1)$ otherwise.

Table 2.3c: Intrinsic parameter redundancies for mark-recovery $\mathrm{T}^J/y/z$ models for a general number of juvenile years J

Model	Model Rank	Parameter Deficiency	Validity $n_1 \geq$	Validity $n_2 \geq$	Method
T^J/C/C	$n_1 - n_2 + 2 + B - G$	0	$J+1$	$J+2$	1
T^J/C/T	$n_1 + B + 1 - G$	0	$J+2$	$2J+2$	1
T^J/C/A	$n_1 + B - G$	1	$J+2$	$2J+2$	2
T^J/C/$\mathrm{A}_{1:J+1}$	$n_1 - n_2 + B + J + 2 - G$	0	$2J$	$2J+1$	1
T^J/C/$\mathrm{A}_{1:J+1}$,T	$\begin{cases} n_1 + n_2 + B - 2J - 1 - G & n_2 < n_1 + J \\ 2n_1 + B - J - 1 - G & n_2 \geq n_1 + J \end{cases}$	$\begin{cases} n_1 - 2n_2 + B + J + 2 - G & n_2 < n_1 + J \\ B - n_2 + 2 - G & n_2 \geq n_1 + J \end{cases}$	$J+1$	$J+1$	2
T^J/C/A,T	E	$n_1 - n_2 + B + 1 - G$	$J+1$	$J+1$	3
T^J/T/C	$n_1 + B - J + 1 - G$	0	$J+1$	$J+2$	1
T^J/T/T	$\begin{cases} n_1 + n_2 + B - 2J - 1 - G & n_2 < n_1 + J \\ 2n_1 + B - J - 1 - G & n_2 \geq n_1 + J \end{cases}$	$\begin{cases} J+1 & n_2 < n_1 + J \\ n_2 - n_1 + 1 & n_2 \geq n_1 + J \end{cases}$	$J+1$	$J+1$	2
T^J/T/A	$n_1 + n_2 + B - J - G$	0	$J+3$	$2J+3$	1
T^J/T/$\mathrm{A}_{1:J+1}$	$n_1 + B + 1 - G$	0	$2J+1$	$2J+2$	1
T^J/T/$\mathrm{A}_{1:J+1}$,T	$\begin{cases} n_1 + n_2 + B - 2J - 1 - G & n_2 < n_1 + J \\ 2n_1 + B - J - 1 - G & n_2 \geq n_1 + J \end{cases}$	$\begin{cases} n_1 - n_2 + B + 1 - G & n_2 < n_1 + J \\ B - J + 1 - G & n_2 \geq n_1 + J \end{cases}$	$J+1$	$J+1$	2
T^J/T/A,T	E	$n_1 + B - J - G$	$J+1$	$J+1$	3
T^J/A/C	$n_1 + B - J + 1 - G$	0	$J+1$	$J+2$	1
T^J/A/T	$n_1 + n_2 + B - J - G$	0	$J+3$	$2J+3$	1
T^J/A/A	$n_1 + B - G$	$n_2 - J$	$J+1$	$2J+1$	2
T^J/A/$\mathrm{A}_{1:J+1}$	$n_1 + B - G$	1	$2J+1$	$2J+2$	2
T^J/A/$\mathrm{A}_{1:J+1}$,T	$\begin{cases} n_1 + 2n_2 + B - 3J - 3 - G & n_2 < n_1 + J \\ 2n_1 + n_2 + B - 2J - 3 - G & n_2 \geq n_1 + J \end{cases}$	$\begin{cases} n_1 - 2n_2 + B + J + 3 - G & n_2 < n_1 + J \\ B - n_2 + 3 - G & n_2 \geq n_1 + J \end{cases}$	$2J+2$	$2J+3$	2
T^J/A/A,T	E	$n_1 + B + J - G$	$J+1$	$J+1$	3
T^J/A,T/z	Identical to A,T/z models: See Table 2.2				

Method 1 = Extension, Method 2 = Reparameterisation, Method 3 = Exhaustive terms;
$E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1; B = \frac{1}{2}J(2n_2 - J + 1)$;
$G = \frac{1}{2}[(n_2 - n_1 - 1)^2 + (n_2 - n_1 - 1)$ when $n_2 \leq n_1 + J$, and $G = \frac{1}{2}(J-1)(2n_2 - J - 2n_1)$ otherwise.

Table 2.4: Estimable parameter combinations for mark-recovery $x^J/y/z$ models

Model	Estimable Parameter Combinations
$\mathrm{C}^J/\mathrm{C}/\mathrm{A}$	$\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$, $\{\prod_{k=1}^{J}\phi_k\phi_a^{i-J+1}(1-\phi_a)\lambda_i\}_{i\,=\,J+1,\ldots,n_2}$
$\mathrm{C}^J/\mathrm{C}/\mathrm{A}_{1:J+1}$	ϕ_a, $\left(\prod_{i=1}^{J}\phi_i\right)\lambda_a$, $\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$
$\mathrm{C}^J/\mathrm{C}/\mathrm{A}_{1:J+1},\mathrm{T}$	ϕ_a, $\{P_{i,i+j-1}\}_{\substack{i\,=\,1,\ldots,n_1-j+1\\ j\,=\,1,\ldots,J}}$, $\{\prod_{k=1}^{J}\phi_k\lambda_{a,i}\}_{i\,=\,J+1,\ldots,n_2}$
$\mathrm{C}^J/\mathrm{T}/\mathrm{T}$	only for $n_2 > n_1+J-1$: $\{\phi_i\}_{i\,=\,1,\ldots,n_1+J-1}$, $\{\lambda_i\}_{i\,=\,1,\ldots,n_1+J-1}$, $\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,n_1+J,\ldots,n_2}}$
$\mathrm{C}^J/\mathrm{T}/\mathrm{A}$	$\{\phi_i\}_{i\,=\,J+1,\ldots,n_2}$, $\{\prod_{k=1}^{J}\phi_k\lambda_i\}_{i\,=\,J+1,\ldots,n_2}$, $\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$
$\mathrm{C}^J/\mathrm{T}/\mathrm{A}_{1:J+1}$	$\{\phi_i\}_{i\,=\,J+1,\ldots,n_2}$, $\prod_{k=1}^{J}\phi_k\lambda_a$, $\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$
$\mathrm{C}^J/\mathrm{T}/\mathrm{A}_{1:J+1},\mathrm{T}$	$\{\phi_i\}_{i\,=\,J+1,\ldots,\min(n_2-1,n_1)}$, $\{P_{i,i+j-1}\}_{\substack{i\,=\,1,\ldots,n_1-j+1\\ j\,=\,1,\ldots,J}}$, $\{\prod_{k=1}^{J}\phi_k\lambda_{a,i}\}_{i\,=\,J+1,\ldots,\min(n_2-1,n_1)}$, $\left\{\prod_{k=1}^{J}\phi_k\prod_{k=\min(n_2,n_1+1)}^{i-1}\phi_k(1-\phi_i)\lambda_{a,i}\right\}_{i\,=\,\min(n_2,n_1+1),\ldots,n_2}$
$\mathrm{C}^J/\mathrm{A},\mathrm{T}/\mathrm{C}$	$\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda\}_{i\,=\,1,\ldots,J}$, $\{P_{i,j}\}_{\substack{i\,=\,J+1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
$\mathrm{C}^J/\mathrm{A},\mathrm{T}/\mathrm{T}$ $[J=1]$	$\{P_{i,j}\}_{\substack{i\,=\,2,\ldots,n_1\\ j\,=\,n_1+1,\ldots,n_2}}$, $\{(1-\phi_1)\lambda_i\}_{i\,=\,1,\ldots,n_1}$, $\left\{\frac{1-P_{i,j}}{(1-\phi_1)\lambda_j}\right\}_{\substack{i\,=\,2,\ldots,n_1\\ j\,=\,i,\ldots,n_1}}$
$\mathrm{C}^J/\mathrm{A},\mathrm{T}/\mathrm{T}$ $[J>1]$	$\{(1-\phi_1)\lambda_i\}_{i\,=\,1,\ldots,n_1}$, $\left\{\frac{\prod_{k=1}^{i}(1-\phi_k)}{1-\phi_1}\right\}_{i\,=\,2,\ldots,J}$, $\left\{\frac{1-P_{i,j}}{(1-\phi_1)\lambda_j}\right\}_{\substack{i\,=\,J+1,\ldots,n_1\\ j\,=\,i,\ldots,n_1}}$
$\mathrm{C}^J/\mathrm{A},\mathrm{T}/\mathrm{A}$	$\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$, $\{\prod_{k=1}^{J}\phi_k\prod_{k=J+1}^{j-1}\phi_{k,k}(1-\phi_{j,j})\lambda_j\}_{j\,=\,J+1,\ldots,n_2}$, $\left\{\frac{P_{i,j}}{P_{1,j-1}}\right\}_{\substack{i\,=\,J+1,\ldots,n_1\\ j\,=\,i,\ldots,n_2}}$
$\mathrm{C}^J/\mathrm{A},\mathrm{T}/\mathrm{A}_{1:J+1}$	$\{\prod_{k=1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,1,\ldots,J}$, $\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,J+1,\ldots,n_2}}$
$\mathrm{T}^J/\mathrm{C}/\mathrm{A}$	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,J\\ j\,=\,i,\ldots,n_2}}$, $\{\lambda_i\}_{i\,=\,1,\ldots,J}$, $\left\{\phi_a^{i-J-1}(1-\phi_a)\lambda_i\right\}_{i\,=\,J+1,\ldots,n_2}$
$\mathrm{T}^J/\mathrm{C}/\mathrm{A}_{1:J+1},\mathrm{T}$	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,i+J-1}}$, $\{\prod_{k=1}^{J}\phi_{k,k}\phi_a^{i-J-1}(1-\phi_a)\lambda_{a,i}\}_{i\,=\,J+1,\ldots,n_2}$, $\left\{\frac{(\prod_{k=1}^{J}\phi_{k,i+k})}{(\prod_{k=1}^{J}\phi_{k,k})\phi_a^i}\right\}_{i\,=\,1,\ldots,J}$
$\mathrm{T}^J/\mathrm{T}/\mathrm{T}$	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,\min(i+J+2,n_2)}}$
$\mathrm{T}^J/\mathrm{T}/\mathrm{A}_{1:J+1},\mathrm{T}$	$\{P_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,i+J-1}}$, $\{P_{1,j}\}_{j\,=\,J+1,\ldots,n_2}$, $\left\{\frac{\prod_{k=1}^{J}\phi_{k,k+j}}{\prod_{k=1}^{J}\phi_{k,k}\prod_{k=1}^{j}\phi_{k+J}}\right\}_{j\,=\,1,\ldots,\min(n_2-J-1,n_1-J)}$
$\mathrm{T}^J/\mathrm{A}/\mathrm{A}$	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,n_1\\ j\,=\,i,\ldots,J}}$, $\{\lambda_i\}_{i\,=\,1,\ldots,J}$, $\{\prod_{k=J+1}^{i-1}\phi_k(1-\phi_i)\lambda_i\}_{i\,=\,J+1,\ldots,n_2}$
$\mathrm{T}^J/\mathrm{A}/\mathrm{A}_{1:J+1}$	$\{\phi_{i,j}\}_{\substack{i\,=\,j,\ldots,J\\ j\,=\,1,\ldots,n_1}}$, $\{\lambda_i\}_{i\,=\,1,\ldots,J}$, $\{\prod_{k=J+1}^{i-1}\phi_k(1-\phi_i)\lambda_a\}_{i\,=\,J+1,\ldots,n_2}$
$\mathrm{T}^J/\mathrm{A}/\mathrm{A}_{1:J+1},\mathrm{T}$	$\{P_{i,j}\}_{\substack{i\,=\,j,\ldots,J\\ j\,=\,1,\ldots,n_1}}$, $\{P_{1,j}\}_{j\,=\,J+1,\ldots,n_2}$, $\left\{\frac{P_{i,i+J}}{P_{1,i+J-1}}\right\}_{i\,=\,2,\ldots,n_2-J}$, $\left\{\frac{\lambda_{a,i}\lambda_{a,i+2}}{\lambda_{a,i+1}^2}\right\}_{i\,=\,J+1,\ldots,\min(n_2-2,n_1-1)}$

Analysis of parameter redundancy is required so that the correct inference can be made from such studies. Determining which mark-recovery models are not intrinsically parameter redundant is important when fitting a candidate model to the data as there may be issues fitting a parameter redundant model to the data observed. These issues may mean that inference for some parameters is not obtained, or potentially estimates are incorrect giving wrong conclusions from the model. We explore parameter redundancy in mark-recovery models in Table 2.2 and see that a large number of complex models are in fact intrinsically parameter redundant. A similar conclusion is obtained from mark-recovery models with different juvenile survival probabilities given in Tables 2.3a, 2.3b and 2.3c. The exact rank or parameter deficiency of a parameter redundant mark-recovery model may also itself be of academic interest. An objective of this work done in this thesis is so a study designer can check our intrinsic parameter redundancy results, and therefore select candidate models so that these models can estimate all of the parameters they require. For y/z mark-recovery models for instance, they may select the T/A mark-recovery model to be a candidate model for their study over a parameter redundant model such as the T/T mark-recovery model. There are other considerations in model selection such as goodness of fit and parameter standard errors, but we believe the parameter redundancy status of a model is also important to consider when selecting a model to use.

We now finish exploring mark-recovery models and use the methods demonstrated in this chapter to explore capture-recapture studies in Chapter 3.

Chapter 3

Capture-Recapture Models

3.1 Capture-recapture background

We now investigate a different ecological model in this chapter by exploring **capture-recapture models**. The purpose of capture-recapture studies is to mark animals at one time point and then to attempt to recapture them alive at future time points, this providing information on animal survival. These marking and recapture occasions can be repeated to give information on survival patterns of animals over long periods of time. Capture-recapture models are different from mark-recovery models as we are recapturing animals alive instead of recovering animals when they have died.

Capture-recapture methods date back as far as Laplace (1786), where they were used to estimate the population size of France, and the first time they were used in an ecological study was in Petersen (1896) to estimate plaice populations. Capture-recapture methods can be distinguished as either being closed or open. In closed capture-recapture models the number of individuals does not change over time, where these changes could be a consequence of either immigration or emigration, or due to the recruitment or death of individuals. Developments in closed capture-recapture models can be traced back to the Schnabel census in Schnabel (1938) and the Petersen-Lincoln and Chapman estimators described in Lincoln (1930) and Chapman (1951) respectively which can be used to provide estimates of animal abundance. We however only explore open capture-recapture models in this chapter as it allows us to mark additional animals over time, i.e. the marked population does not remain constant during the length of the study. A set of models to estimate survival probabilities for this open population case came from Jolly (1965) and Seber (1965) to develop Jolly-Seber models. This was extended in Cormack (1964) to obtain **Cormack-Jolly-Seber models**,

which are the Jolly-Seber models conditional on first capturing of individuals. The importance of these models is due to the model's flexibility as shown by the amount of occasions the model has been used in ecological studies over the last few decades. Review of capture-recapture developments can also be found in Williams et al. (2002) and Amstrup et al. (2005).

Here we review the Cormack-Jolly-Seber model as it is an important development in capture-recapture modelling and links to the capture-recapture models we examine later in this chapter for parameter redundancy. Suppose there are k capture and recapture occasions during a capture-recapture study. There are two types of parameters in the model, survival parameters and recapture parameters. The survival parameters are denoted as ϕ_j which is the probability an animal who is alive at capture occasion j survives until the $(j+1)$th capture occasion, for all $j = 1, \ldots, k$. In capture-recapture studies, the survival probabilities denote an animal's *apparent* survival, where in mark-recovery studies, the survival probabilities denote an animal's *real* survival. The difference between the two terms is that the apparent survival of an animal may be affected by the emigration of that animals from the study area, while the real survival of an animal is not dependent on the animal emigrating or not.

The recapture parameters are denoted as p_j which is the probability that the animal will be recaptured at capture occasion j given the animal is still alive at that point, for all $j = 2, \ldots, k+1$ (we usually denote the first recapture point as $j = 2$ as the animals are marked at the first capture point but cannot be recaptured then). To construct a likelihood we consider the probability a marked animal is not seen alive again after it was last seen at capture occasion j, denoted by χ_j. It can be shown that χ_j satisfies the recurrence equation

$$\frac{1-\chi_j}{\phi_j} = 1 - \chi_{j+1} + \chi_{j+1}p_{j+1}, \tag{3.1}$$

where $\chi_k = 1$. χ is a recursive equation because we are not sure if the animal has survived or not for all the years of the study since it was last recaptured at capture occasion j. The data is summarised in the model with $\mathbf{a}_j$ denoting the number of previously marked animals recaptured at capture occasion j for all $j = 2, \ldots, k+1$, $\mathbf{c}_j$ denoting the number of previously marked animals recaptured for the last time at capture occasion j for all $j = 1, \ldots, k$ (with $\mathbf{c}_1$ denoting when an animal which was marked at the beginning of the study and is not recaptured again in the study) and $\mathbf{v}_j$ denoting the number of previously marked animals known to be alive after capture

occasion j for all $j = 1, \ldots, k$. The likelihood is then given in Cormack (1964) as

$$L = \prod_{j=1}^{k-1} \phi_j^{\mathbf{v}_j} p_{j+1}^{\mathbf{a}_{j+1}} \chi_j^{\mathbf{c}_j} (1 - p_{j+1})^{\mathbf{v}_j - \mathbf{a}_{j+1}}. \tag{3.2}$$

Since the Cormack-Jolly-Seber model was developed in the 1960's, many variants of this model have been widely used. One example of such a variant is the inclusion of age-dependent parameters or parameters dependent on other characteristics such as sex or breeding pattern, such as discussed in Williams et al. (2002). Another variant is to factor in location parameters as they may affect recapture probabilities; the theory behind the model including location parameters is shown in Arnason (1972). Other examples where capture-recapture modelling has been used includes Seber (1986), Pollock et al. (1990), Pollock (2000) and Lindberg (2012). Lebreton et al. (1992) in particular is a widely referenced paper which explores a variety of different capture-recapture models in statistical ecology. These models can be fitted to data using the software packages `MARK` and `SURGE` (which has now been superseded by `M-SURGE` and `E-SURGE`), which are Microsoft Windows programs that provide parameter estimates for a variety of different capture-recapture models including the Cormack-Jolly-Seber model. More detailed information on the programs `MARK` and `SURGE` can be seen in Cooch and White (2014) and Cooch et al. (1997) respectively. A poster by McCrea and Morgan (2008), which was presented at the first ISEC (International Statistical Ecology Conference), shows the many developments in capture-recapture methodology over time and displays the wide range of capture-recapture methods and analysis in use.

An important question that needs to be asked is whether this capture-recapture methodology is still applicable today? J. Andrew Royle stated in Sundvolen, Norway at the third ISEC conference in 2012 that, "Capture-recapture is ubiquitous in ecology - probably the number 1 statistical method." He further mentioned that, "Capture-recapture is more important than ever," showing how important this model still is in modern ecological studies in deriving estimates of abundance and survival probabilities. It is for this reason that we examine parameter redundancy in capture-recapture studies and why it is necessary for these models to continued to be used in practice.

3.2 Capture-recapture models for m-array data

The section considers a certain formulation of the model as a base for further work on capture-recapture modelling. We explored how a d-array can be presented to represent mark-recovery data in Section 2.1. We can now present the data of a capture-recapture study as an **m-array** but we instead have alive recaptures rather than dead recoveries as entries in the array. The Cormack-Jolly-Seber model of Cormack (1964), Jolly (1965) and Seber (1965) can be presented as an m-array if we consider re-marking of animals each time they are captured. This re-marking usually does not involve changing of an animal's mark, instead an animal is just a re-release into the population so the animal can be recaptured again at a future time point.

An example of how data can be presented as an m-array is given in Table 3.1, which is extracted from Lebreton et al. (1992) for data of female Greater Flamingoes (*Phoenicopterus ruber*) between 1982 and 1986. For data presented as an m-array, a number of animals are released at capture occasion i, denoted as R_i, and then possibly recaptured at capture occasion j given they were released at capture occasion i. We do not generally list the animals that were never recaptured in an m-array as the number of animals never recaptured is equal to the number of animals released minus the number of animals we do recapture during the study, but we do it here for illustrative purposes. We can see from the m-array of Table 3.1 that of the 32 animals released at the first capture occasion, 27 animals were recaptured at the first recapture occasion, 4 animals were recaptured at the second recapture occasion, and therefore a single animal was never recaptured at all. We generally re-release animals in capture-recapture studies so that they can observed again, and this happens in this study as the 27 animals that were recaptured at capture occasion 2 were re-released in R_2 so they could be recaptured again. In this case the re-releasing process is simply a way of representing recapturing the animal a multiple number of times during the study and not an actual physical act of re-marking or re-releasing the individuals in the study.

Let the number of capture and recapture occasions in a capture-recapture study to be equal to T. Some care must be taken over the clarity of the model due to the notations used as there actually are $T-1$ different first-capture occasions because there are no new animal captures in the final recapture occasion of the study. Furthermore, there are only $T-1$ recapture occasions as it is not possible to recapture at the first capture occasion in the study as that capture occasion is the first occasion where animals are marked. We continue to let ϕ_j which is the probability an animal who is alive

Table 3.1: An m-array representation of data from a capture-recapture study of female Greater Flamingoes from Lebreton et al. (1992)

Year of release	Number of flamingoes released	Year of next recapture					Never recaptured
		1983	1984	1985	1986	1987	
1982	$R_1 = 32$	27	4	0	0	0	1
1983	$R_2 = 80$		57	7	3	4	9
1984	$R_3 = 112$			58	18	4	32
1985	$R_4 = 132$				109	13	10
1986	$R_5 = 211$					169	42

at capture occasion j survives until the $(j+1)$th capture occasion, for all $j = 1, \ldots, T$, and p_j denote the probability that an animal is recaptured alive at capture occasion j, for all $j = 2, \ldots, T+1$ (i.e. $j = 1$ would represent a capture occasion of only first-captures and no possible recaptures are not possible). We demonstrate what the probability of a recapture possibility of an animal is in Example 3.1.

Example 3.1: Consider an animal which has been marked and released at the second capture occasion $i = 2$, and is next recaptured at the fifth capture occasion in the study $j = 5$. We know that the animal has survived until the fifth capture occasion during the study, while the animal was not recaptured during the third and fourth capture occasions during the study. The probability of this, denoted as $P_{2,5}$, is given by

$$P_{2,5} = \phi_2(1 - p_3)\phi_3(1 - p_4)\phi_4 p_5.$$

□

We can show that the probability an animal is marked at capture occasion i and is next recaptured at recapture occasion j, as denoted by $P_{i,j}$, can be expressed as

$$P_{i,j} = \left(\prod_{m=i}^{j-1} \phi_m\right)\left(\prod_{m=i+1}^{j-1} (1 - p_m)\right) p_j, \tag{3.3}$$

where $\prod_{m=i+1}^{i}(1 - p_m) = 1$ for the case that an animal is recaptured at the first recapture occasion after first-capture. All of the possible capture-recapture probabilities can be expressed as the model's p-array in the matrix $\mathbf{P}$, which has the terms $P_{i,j}$ for the diagonal and upper diagonal terms where $i \geq j$ and 0 otherwise, such as for $T = 3$

in the matrix below.

$$\mathbf{P} = \begin{bmatrix} P_{1,2} & P_{1,3} & P_{1,4} \\ 0 & P_{2,3} & P_{2,4} \\ 0 & 0 & P_{3,4} \end{bmatrix} = \begin{bmatrix} \phi_1 p_2 & \phi_1\phi_2(1-p_2)p_3 & \phi_1\phi_2\phi_3(1-p_2)(1-p_3)p_4 \\ 0 & \phi_2 p_3 & \phi_2\phi_3(1-p_3)p_4 \\ 0 & 0 & \phi_3 p_4 \end{bmatrix}.$$

The probability of a captured animal not being recaptured at all during the study is given as

$$Pr(\text{A marked animal is not recaptured in the study}) = \left(1 - \sum_{j=i+1}^{T+1} P_{i,j}\right),$$

for any release occasion i. If we let $N_{i,j}$ denote the number of animals released at capture occasion i, for all $i = 1, \ldots, T$ and recaptured at recapture occasion j for all $j = 2, \ldots, T+1$, and further let R_i denote the number of animals captured at capture occasion i, for all $i = 1, \ldots, T$, then in a similar way to the mark-recovery model in Equation (2.2), the likelihood can be expressed as

$$L = \prod_{i=1}^{T} \left\{ \prod_{j=i+1}^{T+1} P_{i,j}^{N_{i,j}} \left(1 - \sum_{j=i+1}^{T+1} P_{i,j}\right)^{R_i - \sum_{j=i+1}^{T+1} N_{i,j}} \right\}. \tag{3.4}$$

Catchpole and Morgan (1997) show that we only need to consider the $P_{i,j}$ terms as an exhaustive summary as long as all recaptures occur, i.e. $N_{i,j} > 0$ for all i and j. This means we do not have to consider the terms where the animals are not recaptured during the study in an exhaustive summary.

We note that this likelihood in Equation (3.4) is the same as the previous Cormack-Jolly-Seber likelihood in Equation (3.2). This can be seen as the $\prod_{i=1}^{T-1}\left(1 - \sum_{j=i+1}^{T+1} P_{i,j}\right)$ term in Equation (3.4) is equivalent to $\prod_{j=1}^{T-1} \chi_j^{\mathbf{c}_j}$ term in Equation (3.2) as this denotes the number of animals where are never recaptured, and the remaining terms in Equation (3.4) are equivalent to the $\prod_{j=i}^{T} P_{i,j}$ term in Equation (3.2) as they are equal to the possible capture-histories that can be observed.

We can also vary the parameter dependencies in this model by considering constant parameters instead of time-dependent ones for both survival and recapture parameters. We follow a similar notation to mark-recovery models and denote the model as y/z where y refers to the survival parameters and z refers to the recapture parameters, and these can be either constant (C) or time-dependent (T). The Cormack-Jolly-Seber

model is the model where both survival and recapture parameters are time-dependent, and we show this in Example 3.2 below.

Example 3.2 - The m-array capture-recapture T/T model: [See electronic appendix `example3.2.mw`] The m-array with a total of three capture and recapture occasions, with time-dependent survival and recapture probabilities, can be generated by the `Maple` code

```
P := capmodmarray(2,2,3);
# Inputs: (y,z,T);
# y = survival probability; z = recapture probability;
# for y and z: '1'=constant probabilities and '2'=time-dependent;
# T = number of capture and recapture occasions in the study.
```

`Maple` returns the p-array probabilities as

$$\mathbf{P} = \begin{bmatrix} \phi_1 p_2 & \phi_1\phi_2(1-p_2)p_3 & \phi_1\phi_2\phi_3(1-p_2)(1-p_3)p_4 \\ 0 & \phi_2 p_3 & \phi_2\phi_3(1-p_3)p_4 \\ 0 & 0 & \phi_3 p_4 \end{bmatrix}.$$

From this p-array, we can then find the parameter deficiency of the model as described in Section 2.2:

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta);
r := Rank(DD); d := Dimension(theta)-r;
```

The exhaustive summary is

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_1 p_2 \\ \phi_1\phi_2(1-p_2)p_3 \\ \phi_1\phi_2\phi_3(1-p_2)(1-p_3)p_4 \\ \phi_2 p_3 \\ \phi_2\phi_3(1-p_3)p_4 \\ \phi_3 p_4 \end{bmatrix},$$

with parameters $\boldsymbol{\theta} = [p_2, p_3, p_4, \phi_1, \phi_2, \phi_3]^T$ and derivative matrix

$$\mathbf{DD} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} \phi_1 & -\phi_1\phi_2 p_3 & -\phi_1\phi_2\phi_3(1-p_3)p_4 & 0 & 0 & 0 \\ 0 & \phi_1\phi_2(1-p_2) & -\phi_1\phi_2\phi_3(1-p_2)p_4 & \phi_2 & -\phi_2\phi_3 p_4 & 0 \\ 0 & 0 & \phi_1\phi_2\phi_3(1-p_2)(1-p_3) & 0 & \phi_2\phi_3(1-p_3) & \phi_3 \\ p_2 & \phi_2(1-p_2)p_3 & \phi_2\phi_3(1-p_2)(1-p_3)p_4 & 0 & 0 & 0 \\ 0 & \phi_1(1-p_2)p_3 & \phi_1\phi_3(1-p_2)(1-p_3)p_4 & p_3 & \phi_3(1-p_3)p_4 & 0 \\ 0 & 0 & \phi_1\phi_2(1-p_2)(1-p_3)p_4 & 0 & \phi_2(1-p_3)p_4 & p_4 \end{bmatrix}.$$

The model rank and parameter deficiency of the model are equal to

$$\mathbf{r} := \mathbf{5}, \mathbf{d} := \mathbf{1}.$$

This shows that in this case the T/T capture-recapture model has 5 estimable parameter combinations and a parameter deficiency of 1. □

We obtain the following general intrinsic results for capture-recapture models:

- The C/C model which has both constant survival and recapture probabilities is not parameter redundant and we can estimate both parameters ϕ and p. This can be proved for a general T trivially by using the extension theorem of Theorem 2.2.

- The C/T and T/C models are also not parameter redundant and all the $T+1$ different parameters in the model can be estimated in both cases. This can also be proved for a general T by the extension theorem of Theorem 2.2.

- The T/T model has a parameter deficiency of 1 with $2T-1$ estimable parameter combinations in the model. The proof of this result is by using the reparameterisation theorem of Theorem 2.3. The original parameters ϕ_j and p_{j+1} are estimable for $j = 1, \ldots, T-1$, but the parameters ϕ_T and p_{T+1} are confounded as the product $\phi_T p_{T+1}$. This result is well known as this time-dependent case is a representation of the Cormack-Jolly-Seber which has this confounding of the final survival and recapture parameters, e.g. as shown in Lebreton et al. (1992, Table 3) where they can only estimate ϕ_1, p_2 and $\phi_2 p_3$ for a $T = 2$ study.

We now explore a link between these results and the mark-recovery results of Table 2.2 using a transformation in the next section.

3.3 A transformation between capture-recapture models for m-array data and mark-recovery models for d-array data

When performing the analyses of Sections 2.9 and 3.2, we notice the strong links between the results for y/z capture-recapture models and Table 2.2 for y/z mark-recovery models. If you put both sets of results along side each other, you see that the capture-recapture y/z model has the exactly the same model rank and parameter deficiency as the z/y mark-recovery model. This suggests that there is a reparameterisation which equates mark-recovery and m-array capture-recapture models. Lebreton et al. (1995) notice this link between capture-recapture and mark-recovery models and the appropriate transformation is listed in the appendix of their paper. Below we demonstrate how a reparameterisation relates between the two models, and then show in Example 3.3 how such a transformation can be performed.

- Transforming a capture-recapture y/C model to a mark-recovery C/y model: Let $p = (1 - \phi)$.
- Transforming a capture-recapture y/T model to a mark-recovery T/y model: Let $p_{i+1} = (1 - \phi_i)$ for all $i = 1, \ldots, k$.
- Transforming a capture-recapture C/z model to a mark-recovery z/C model: Let $\phi = \lambda$ for when the animal is recaptured during the capture occasion immediately after the animal has been captured. Further let $\phi = 1$ for when the animal is recaptured at any future capture occasion. We can then obtain the result that the iteration $\lambda = 1 \times 1 \times \cdots \times \phi^*$ applies for when the animal is recaptured at capture occasion j given that it was first-captured at capture occasion i.
- Transforming a capture-recapture T/z model to a mark-recovery z/T model: Let $\phi_i^* = \lambda_i$ for when the animal is recaptured during the capture occasion immediately after the animal has been captured. Further let $\phi_j^{'} = \lambda_j/\lambda_{j-1}$ for when the animal is recaptured at any future capture occasion. We can then obtain the result that the iteration $\lambda_j = \phi_j^{'}\phi_{j-1}^{'}\phi_{j-2}^{'}\ldots\,\phi_{i+2}^{'}\phi_{i+1}^{'}\phi_i^*$ applies for when the animal is recaptured at capture occasion j given that it was first-captured at capture occasion i.

Example 3.3: This reparameterisation can used for example to convert the capture-recapture C/T model into a mark-recovery T/C model for $T = 3$. By applying the

transformation $p_i = (1 - \phi_{i-1})$ and $(1 - p_i) = \phi_{i-1}$ for all $i = 2, \ldots, T + 1$, as well the transformation where $\phi = \lambda$ for the first ϕ in an expression and then $\phi = 1$ for the remaining ϕ's in the expression (e.g. $\phi^3 = \lambda \times 1 \times 1 = \lambda$) , we get

$$\begin{aligned} \mathbf{P} &= \begin{bmatrix} \phi p_2 & \phi^2(1-p_2)p_3 & \phi^3(1-p_2)(1-p_3)p_4 \\ 0 & \phi p_3 & \phi^2(1-p_3)p_4 \\ 0 & 0 & \phi p_4 \end{bmatrix} \\ &\rightarrow \begin{bmatrix} (1-\phi_1)\lambda & \phi_1(1-\phi_2)\lambda & \phi_1\phi_2(1-\phi_3)\lambda \\ 0 & (1-\phi_2)\lambda & \phi_2(1-\phi_3)\lambda \\ 0 & 0 & (1-\phi_3)\lambda \end{bmatrix}. \end{aligned}$$

This shows how a capture-recapture model can be transformed into a mark-recovery model. □

This transformation is useful as a check as by the use of the reparameterisation theorem of Theorem 2.3 we can deduce capture-recapture results from mark-recovery results. It is also useful to show that the mark-recovery and capture-recapture models have similarities between each other.

3.4 The capture-history formulation of capture-recapture models and a simpler exhaustive summary

We now present an alternative formulation of capture-recapture models by listing the individual **capture-histories** as a way of forming an exhaustive summary. A capture-history is a way of presenting the recapture data from a single animal in the study. This formulation using these capture-histories has the advantage of allowing the consideration of age-dependent parameters, something which is not possible in the m-array formulation of the capture-recapture model. It also allows us to more conveniently explore extrinsic parameter redundancy in Section 3.7.

Let n_1 denote the number of first-capture occasions in the study and n_2 denote the number of recapture occasions in the study, where $n_1 \leq n_2$ as there at least is as many recapture occasions as there are first-capture occasions. We mark the animals at age 0 if we consider an age-dependent model, and therefore do not consider the animals that may live to a greater age than n_2 and are alive at the conclusion of the study. Frequently in capture-recapture studies, there are same number of T capture and recapture occasions so that $n_1 = n_2 = T - 1$, but it is possible to have a study

where $n_2 > n_1$. These capture and recapture occasions are usually annual but other periods are possible. To allow age-dependency in our parameters we let $\phi_{i,j}$ represent the probability that an animal of age $i-1$ at time j survives until time $j+1$, given that the animal has survived up to that point since it was first-captured for $i = 1, \ldots, n_2$ and $j = i, \ldots, n_2$. We further let $p_{i,j}$ denote the probability that an animal of age $i-1$ is recaptured at capture occasion j (i.e. recapture occasion $j-1$) for $i = 2, \ldots, n_2+1$ and $j = i, \ldots, n_2+1$.

The capture-history of an animal is a result of repeated observations recording whether the animal was recaptured at each capture occasion. The notation that has historically been used for capture-recapture data is the use of binary code to represent these capture-histories. In a single capture-history, a '1' indicates when the animal was recaptured, and a '0' indicates when the animal was *not* recaptured. A '1' is also used to represent the capture occasion when it was first-captured. Examples 3.4 and 3.5 below show how we can represent the probabilities of single capture-histories using survival and recapture probabilities.

Example 3.4: The capture-history 1001011 is observed for a study where there are $T = 7$ capture occasions. This indicates that the animal was first captured at the beginning of the study, was not recaptured at capture occasions two, three and five, but was recaptured at capture occasions four, six and seven. If we assume that survival and recapture probabilities are known to be only time-dependent, this capture-history has the probability

$$Pr(1001011) = \phi_1(1-p_2)\phi_2(1-p_3)\phi_3 p_4 \phi_4(1-p_5)\phi_5 p_6 \phi_6 p_7.$$

□

Example 3.5: An alternative capture-history is 0100110 where there are still $T = 7$ capture occasions. This indicates that the animal was first captured at the second capture occasion with there still being five possible occasions of recapture. If we assume that survival and recapture probabilities are known to be only time-dependent, this capture-history has the probability

$$Pr(0100110) = \phi_2(1-p_3)\phi_3(1-p_4)\phi_4 p_5 \phi_5 p_6 \chi_6,$$

where $\chi_6 = (1-\phi_6) + \phi_6(1-p_7) = 1 - \phi_6 p_7$. The term χ_6 represents the probability of not being recaptured at any further capture occasions after the sixth capture oc-

casion, and gives the probability that the animal either did not survive up until the seventh capture occasion given the animal was alive at the sixth capture occasion, or the animal actually did survive but was not recaptured at capture occasion seven. □

The probability of a particular capture-history can be generalised as given in Hubbard et al. (2014). Define $\chi_{i,j}$ as the probability an animal released at capture occasion i is not recaptured again since capture occasion j to be

$$\chi_{i,j} = (1 - \phi_{i,j}) + \phi_{i,j}\,(1 - p_{i+1,j+1})\,\chi_{i+1,j+1}, \tag{3.5}$$

with $\chi_{i,n_2} = 1$ for all i. If we suppose an animal was first captured at capture occasion a and was last recaptured at capture occasion b, then the probability of a particular capture-history, h_X, is given by

$$Pr(h_X) = \left[\prod_{k=a+1}^{b} \phi_{k-a,k-1}\left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\}\right]\chi_{b-a+1,b}, \tag{3.6}$$

where δ_k denotes the value taken by the capture-history at time k. If we let N be the number of distinct capture-histories observed, then the likelihood is given by

$$L = \prod_{X=1}^{N} Pr(h_X). \tag{3.7}$$

A suitable exhaustive summary for the model can be generated containing all the probabilities of the distinct capture-histories as separate exhaustive summary terms. Table 3.2a shows all of the possible capture-histories for $n_1 = n_2 = 1, 2, 3, 4$, where the probabilities of each capture-history are listed in Table 3.2b where it is assumed that survival and recapture probabilities are both only time-dependent. Observe that it is clear from Tables 3.2a and 3.2b that the number of possible capture-histories grows large very quickly.

By using the formula for the sum of a geometric series, it can shown that there are $\sum_{i=n_2-n_1+1}^{n_2} 2^i = 2^{n_2+1} - 2^{n_2-n_1+1}$ possible combinations of capture-history in this capture-recapture model. For a large number of capture occasions, this raises a problem as `Maple` could encounter memory limitations. A simpler exhaustive summary of the model is therefore required due to the large number of exhaustive summary terms in the original model, so that we can generate general intrinsic parameter redundancy results for capture-recapture models. We can then use the result from Cole et al.

Table 3.2a: All the possible capture-histories in capture-recapture models in binary notation

$T = 2$: $n_1 = n_2 = 1$	$T = 3$: $n_1 = n_2 = 2$	$T = 4$: $n_1 = n_2 = 3$	$T = 5$: $n_1 = n_2 = 4$
11	011	0011	00011
10	111	0111	00111
	101	0101	00101
	010	1111	01111
	110	1011	01011
	100	1101	01101
		1001	01001
		0010	11111
		0110	10111
		0100	11011
		1110	11101
		1010	10011
		1100	10101
		1000	11001
			10001
			00010
			00110
			01110
			01010
			00100
			01100
			01000
			11110
			10110
			11010
			11100
			10010
			10100
			11000
			10000

(2010) that a reparameterised version of the model will still have the same parameter deficiency of the original model as shown in Section 2.3.

In order to create a simpler exhaustive summary we use a method of proof specifically designed to prove a reparameterisation which generates a sufficient exhaustive summary. This is called the **two-stage extension theorem** and develops the standard extension theorem of Theorem 2.2 so we can find exhaustive summaries where the exhaustive summary terms are simpler. Theorem 3.1 has been used in Cole and Morgan (2010a), Cole (2012) and Hubbard et al. (2014).

In the standard extension theorem of Catchpole and Morgan (1997, Theorem 6) and Cole et al. (2010, Theorem 3), we begin with exhaustive summary $\boldsymbol{\kappa}_1^S(\boldsymbol{\theta}_1)$ (where we denote S for the exhaustive summary using the standard extension theorem) which has parameters $\boldsymbol{\theta}_1$. Then we extend this model to give the exhaustive summary

Table 3.2b: All the possible capture-history probabilities corresponding to Table 3.2a

$T=2$: $n_1=n_2=1$	$T=3$: $n_1=n_2=2$	$T=4$: $n_1=n_2=3$	$T=5$: $n_1=n_2=4$
$\phi_1 p_2$	$\phi_2 p_3$	$\phi_3 p_4$	$\phi_4 p_5$
χ_1	$\phi_1 p_2 \phi_2 p_3$	$\phi_2 p_3 \phi_3 p_4$	$\phi_3 p_4 \phi_4 p_5$
	$\phi_1(1-p_2)\phi_2 p_3$	$\phi_2(1-p_3)\phi_3 p_4$	$\phi_3(1-p_4)\phi_4 p_5$
	χ_2	$\phi_1 p_2 \phi_2 p_3 \phi_3 p_4$	$\phi_2 p_3 \phi_3 p_4 \phi_4 p_5$
	$\phi_1 p_2 \chi_2$	$\phi_1(1-p_2)\phi_2 p_3 \phi_3 p_4$	$\phi_2(1-p_3)\phi_3 p_4 \phi_4 p_5$
	χ_1	$\phi_1 p_2 \phi_2(1-p_3)\phi_3 p_4$	$\phi_2 p_3 \phi_3(1-p_4)\phi_4 p_5$
		$\phi_1(1-p_2)\phi_2(1-p_3)\phi_3 p_4$	$\phi_2(1-p_3)\phi_3(1-p_4)\phi_4 p_5$
		χ_3	$\phi_1 p_2 \phi_2 p_3 \phi_3 p_4 \phi_4 p_5$
		$\phi_2 p_3 \chi_3$	$\phi_1(1-p_2)\phi_2 p_3 \phi_3 p_4 \phi_4 p_5$
		χ_2	$\phi_1 p_2 \phi_2(1-p_3)\phi_3 p_4 \phi_4 p_5$
		$\phi_1 p_2 \phi_2 p_3 \chi_3$	$\phi_1 p_2 \phi_2 p_3 \phi_3(1-p_4)\phi_4 p_5$
		$\phi_1(1-p_2)\phi_2 p_3 \chi_3$	$\phi_1(1-p_2)\phi_2(1-p_3)\phi_3 p_4 \phi_4 p_5$
		$\phi_1 p_2 \chi_2$	$\phi_1(1-p_2)\phi_2 p_3 \phi_3(1-p_4)\phi_4 p_5$
		χ_1	$\phi_1 p_2 \phi_2(1-p_3)\phi_3(1-p_4)\phi_4 p_5$
			$\phi_1(1-p_2)\phi_2(1-p_3)\phi_3(1-p_4)\phi_4 p_5$
			χ_4
			$\phi_3 p_4 \chi_4$
			χ_3
			$\phi_2 p_3 \phi_3 p_4 \chi_4$
			$\phi_2(1-p_3)\phi_3 p_4 \chi_4$
			$\phi_2 p_3 \phi_3(1-p_4)\chi_4$
			χ_2
			$\phi_1 p_2 \phi_2 p_3 \phi_3 p_4 \chi_4$
			$\phi_1(1-p_2)\phi_2 p_3 \phi_3 p_4 \chi_4$
			$\phi_1 p_2 \phi_2(1-p_3)\phi_3 p_4 \chi_4$
			$\phi_1 p_2 \phi_2 p_3 \chi_3$
			$\phi_1(1-p_2)\phi_2(1-p_3)\phi_3 p_4 \chi_4$
			$\phi_1(1-p_2)\phi_2 p_3 \chi_3$
			$\phi_1 p_2 \chi_2$
			χ_1
$\chi_j = (1-\phi_j) + \phi_j(1-p_{j+1})\chi_{j+1}$ with $\chi_{n_2+1}=1$ for the time-dependent case.			

$\boldsymbol{\kappa}^S(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = [\boldsymbol{\kappa}_1^S(\boldsymbol{\theta}_1),\ \boldsymbol{\kappa}_2^S(\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2)]^T$ with parameters $\boldsymbol{\theta}^S = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T$. If $\partial \boldsymbol{\kappa}_1^S(\boldsymbol{\theta}_1)/\partial \boldsymbol{\theta}_1$ is full rank and $\partial \boldsymbol{\kappa}_2^S(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)/\partial \boldsymbol{\theta}_2$ is full rank, then $\partial \boldsymbol{\kappa}^S(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)/\partial \boldsymbol{\theta}^S$ is also full rank.

In the two-stage extension theorem, we begin with exhaustive summary $\boldsymbol{\kappa}_1^O(\boldsymbol{\theta}_1)$, which can be partitioned as $\boldsymbol{\kappa}_1^O(\boldsymbol{\theta}_1)$=$[\boldsymbol{\kappa}_1^E(\boldsymbol{\theta}_{1,1}),\ \boldsymbol{\kappa}_2^E(\boldsymbol{\theta}_{1,1},\ \boldsymbol{\theta}_{1,2})]$ with parameters $\boldsymbol{\theta}_1 = [\boldsymbol{\theta}_{1,1},\ \boldsymbol{\theta}_{1,2}]^T$. This exhaustive summary is then extended to $\boldsymbol{\kappa}_2^O(\boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2}) = [\boldsymbol{\kappa}_1^E(\boldsymbol{\theta}_{1,1}), \boldsymbol{\kappa}_2^E(\boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2})]$, with parameters $\boldsymbol{\theta}_2 = [\boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2}]^T$. Let $\boldsymbol{\theta}_2^E = [\boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2}]^T$. Then we have the two-stage extension theorem of Theorem 3.1 below.

Theorem 3.1. *If $\partial \boldsymbol{\kappa}_1^O(\boldsymbol{\theta}_1)/\partial \boldsymbol{\theta}_1$, $\partial \boldsymbol{\kappa}_1^E(\boldsymbol{\theta}_{1,1})/\partial \boldsymbol{\theta}_{1,1}$, and $\partial \boldsymbol{\kappa}_2^E(\boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2})/\partial \boldsymbol{\theta}_2^E$ are all full rank, then $\boldsymbol{\kappa}_2^O(\boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2})$ is also full rank.*

The proof of Theorem 3.1 follows the same form as the standard extension theorem

as stated in Hubbard et al. (2014). It can also be derived indirectly from Meyer (1973, Theorem 4.2). This two-stage extension theorem is then used in Theorem 3.2 to prove there is a simpler exhaustive summary than the exhaustive summary containing all the capture-histories as individual exhaustive summary terms.

Theorem 3.2. *A simpler exhaustive summary for the capture-recapture model consists of the terms:*

- $s_{i,j} = \phi_{i,j} p_{i+1,j+1}$ *for all* $i = 1, \ldots, n_2$ *and* $j = i, \ldots, \min(n_1 + i - 1, n_2)$,
- *and* $t_{i,j} = \phi_{i,j}(1 - p_{i+1,j+1})$ *for all* $i = 1, \ldots, n_2 - 1$ *and* $j = i, \ldots, \min(n_1 + i - 1, n_2 - 1)$.

The proof of Theorem 3.2 can be found in Appendix B.1 as well as in the supplementary material of Hubbard et al. (2014). Theorem 3.2 is also given in Hubbard et al. (2014, Theorem 1a.). We can code a `Maple` procedure in which all of the **s** and **t** terms are generated using the simpler exhaustive summary of Theorem 3.2. These **s** and **t** terms can then be used to examine the intrinsic parameter redundancy of capture-recapture models. This also allows us to have constant (C), time-dependent (T), age-dependent (A) and age- and time-dependent (A,T) options on both survival and recapture parameters in the model. The `Maple` procedure for this can be seen in Example 3.6 below.

Example 3.6 - The capture-recapture A,T/A,T intrinsic model: [See electronic appendix `example3.6.mw`] We can generate the simpler exhaustive summary terms using Theorem 3.2 to obtain the intrinsic parameter deficiency for capture-recapture models. We illustrate this in a capture-recapture model with four capture and recapture occasions which has both age- and time-dependency on the survival and recapture parameters. Observe that we demonstrate this `Maple` procedure `capmodintrinsic` by using an example which assumes age- and time-dependence is valid for the data, and this assumption may not be applicable for all capture-recapture studies.

```
P := capmodintrinsic(4,4,4,4);
# Inputs: (y,z,n1,n2);
# y = survival probability; z = recapture probability;
# for y and z: '1'=C, '2'=T, '3'=A, '4'=A,T;
# n1 = number of first-capture occasions;
# n2 = number of recapture occasions.
```

`Maple` returns the matrix

$$\mathbf{P} = \begin{bmatrix} s_{1,4} & s_{1,3} & s_{1,2} & s_{1,1} \\ 0 & s_{2,4} & s_{2,3} & s_{2,2} \\ 0 & 0 & s_{3,4} & s_{3,3} \\ 0 & 0 & 0 & s_{4,4} \\ 0 & t_{1,3} & t_{1,2} & t_{1,1} \\ 0 & 0 & t_{2,3} & t_{2,2} \\ 0 & 0 & 0 & t_{3,3} \end{bmatrix} = \begin{bmatrix} \phi_{1,4}p_{2,5} & \phi_{1,3}p_{2,4} & \phi_{1,2}p_{2,3} & \phi_{1,1}p_{2,2} \\ 0 & \phi_{2,4}p_{3,5} & \phi_{2,3}p_{3,4} & \phi_{2,2}p_{3,3} \\ 0 & 0 & \phi_{3,4}p_{4,5} & \phi_{3,3}p_{4,4} \\ 0 & 0 & 0 & \phi_{4,4}p_{5,5} \\ 0 & \phi_{1,3}(1-p_{2,4}) & \phi_{1,2}(1-p_{2,3}) & \phi_{1,1}(1-p_{2,2}) \\ 0 & 0 & \phi_{2,3}(1-p_{3,4}) & \phi_{2,2}(1-p_{3,3}) \\ 0 & 0 & 0 & \phi_{3,3}(1-p_{4,4}) \end{bmatrix}.$$

The **s** and **t** terms are represented in the matrix form above for convenience, where $\boldsymbol{\kappa}$ is given as the vector of all these non-zero terms in **P**. The parameter deficiency of the model can then be found as described in Section 2.2.

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta):
r := Rank(DD); d := Dimension(theta)-r;
```

The exhaustive summary is

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_{1,4}p_{2,5} \\ \phi_{1,3}p_{2,4} \\ \phi_{1,2}p_{2,3} \\ \phi_{1,1}p_{2,2} \\ \phi_{2,4}p_{3,5} \\ \vdots \\ \phi_{2,2}(1-p_{3,3}) \\ \phi_{3,3}(1-p_{4,4}) \end{bmatrix},$$

with parameters $\boldsymbol{\theta} = [\phi_{1,1},\ \phi_{1,2},\ \phi_{1,3},\ \phi_{1,4},\ \phi_{2,2},\ \phi_{2,3},\ \phi_{2,4},\ \phi_{3,3},\ \phi_{3,4},\ \phi_{4,4},\ p_{2,2},\ p_{2,3},\ p_{2,4},\ p_{2,5},\ p_{3,3},\ p_{3,4},\ p_{3,5},\ p_{4,4},\ p_{4,5},\ p_{5,5}]^T$. The derivative matrix is given in the `Maple` file `example3.6.mw` only to save space. This derivative matrix has the model rank and parameter deficiency

$\mathbf{r} := \mathbf{16}, \mathbf{d} := \mathbf{4}$.

The A,T/A,T model with four first-capture occasions and four recapture occasions has 16 estimable parameter combinations and a parameter deficiency of 4. It can

be seen in `Maple` that the confounded parameters in the model are $\phi_{1,4}p_{2,5}$, $\phi_{2,4}p_{3,5}$, $\phi_{3,4}p_{4,5}$ and $\phi_{4,4}p_{5,5}$ with every other parameter being in principle estimable. □

We now generate general intrinsic parameter redundancy results listed in the next section.

3.5 Intrinsic capture-recapture y/z model results

If we use the exhaustive summary of Theorem 3.2 along with the standard extension and reparameterisation theorems as described in Section 2.3, we can find general intrinsic parameter redundancy results for capture-recapture models with any combination of age- and/or time-dependent parameters. This is shown in Table 3.3 where the second column gives the rank of the derivative matrix, which is equal to the number of estimable parameters in the model, and the third column refers to the parameter deficiency where a model is parameter redundant when $d > 0$. There are n_1 first-capture occasions and n_2 recapture occasions, where $n_2 \geq 2$ with $n_2 \geq n_1$. The final column refers to the method of proof used as described in Section 2.3.

Table 3.4 gives the estimable parameter combinations for each model, which uses the method of solving a system of Lagrange partial differentiation equations as detailed in Section 2.2. Any model using the reparameterisation theorem proof is based on the reparameterisations given in Table 3.4. The results of Tables 3.3 and 3.4 assume there are perfect data observed with at least one observation of each possible distinct capture-history.

Table 3.3 shows that a high number of y/z capture-recapture models are not intrinsically parameter redundant. Furthermore, for all those that have a small parameter deficiency, Table 3.4 shows that *all* the parameters can still be estimated apart from the *last set* of survival and recapture parameters, i.e all the ϕ_i and p_{i+1} parameters can be estimated for $i = 1, \ldots, n_2 - 1$. This is worth emphasising as even though some of the models are parameter redundant, nearly all of the parameters can still be estimated in the model in theory.

Table 3.3: Intrinsic parameter redundancies for capture-recapture y/z models

Model	Rank	Deficiency	Method of Proof
C/C	2	0	Extension Theorem
C/T	n_2+1	0	Extension Theorem
C/A	n_2+1	0	Extension Theorem
C/A,T	$E+1$	0	Extension Theorem
T/C	n_2+1	0	Extension Theorem
T/T	$2n_2-1$	1	Reparameterisation
T/A	$2n_2$	0	Extension Theorem
T/A,T	$E+n_2-1$	1	Reparameterisation
A/C	n_2+1	0	Extension Theorem
A/T	$2n_2$	0	Extension Theorem
A/A	$2n_2-1$	1	Reparameterisation
A/A,T	$E+n_2-1$	1	Reparameterisation
A,T/C	$E+1$	0	Extension Theorem
A,T/T	$E+n_2-1$	1	Reparameterisation
A,T/A	$E+n_2-1$	1	Reparameterisation
A,T/A,T	$2E-n_1$	n_1	Reparameterisation
$E = n_1n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.			

3.6 Capture-recapture models with juvenile survival probabilities

A natural extension to the capture-recapture model is one which has different survival probabilities for when the animal is a juvenile. We viewed similar analyses to these for mark-recovery models with a set of juvenile survival probabilities in Section 2.10. Such capture-recapture examples where a differentiation between juvenile and adult survival probabilities has been made includes data on grey seals in Schwarz and Stobo (2000) and on thrush in California in Gardalia et al. (2003).

Capture-recapture models which have juvenile survival probabilities where the animal is a juvenile for J recapture occasions (for $1 \leq J < n_2 - 1$) are denoted as $x^J/y/z$, where x^J denotes whether the juvenile survival probabilities are constant or time-dependent for J recapture occasions, y denotes the adult survival probabilities and z denotes the recapture probabilities. If x is constant the model has the survival parameters $\phi_1, \phi_2, \ldots, \phi_J$ for each different occasion of first-capture i, and if x is time-dependent, then the model has the survival parameters $\phi_{i,1}, \phi_{i,2}, \ldots, \phi_{i,J}$ for each

Table 3.4: Estimable parameter combinations for capture-recapture y/z models

Model	Estimable Parameter Combinations
C/C	ϕ, p
C/T	ϕ, $\{p_i\}_{i\,=\,2,\ldots,\,n_2\,+\,1}$
C/A	ϕ, $\{p_i\}_{i\,=\,2,\ldots,\,n_2\,+\,1}$
C/A,T	ϕ, $\{p_{i,j}\}_{\substack{i\,=\,2,\ldots,\,n_2\,+\,1\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2\,+\,1)}}$
T/C	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2}$, p
T/T	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2\,-\,1}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2}$, $\phi_{n_2}p_{n_2+1}$
T/A	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2\,+\,1}$
T/A,T	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2\,-\,1}$, $\{p_{i,j}\}_{\substack{i\,=\,2,\ldots,\,n_2\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2)}}$, $\{\phi_{n_2}p_{i+1,n_2+1}\}_{i\,=\,n_2\,-\,n_1\,+\,1,\ldots,\,n_2}$
A/C	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2}$, p
A/T	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2\,+\,1}$
A/A	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2\,-\,1}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2}$, $\phi_{n_2}p_{n_2+1}$
A/A,T	$\{\phi_i\}_{i\,=\,1,\ldots,\,n_2\,-\,1}$, $\phi_{n_2}p_{n_2+1,n_2+1}$, $\{p_{i,j}\}_{\substack{i\,=\,2,\ldots,\,n_2\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2)}}$
A,T/C	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,\,n_2\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2)}}$, p
A,T/T	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,\,n_2\,-\,1\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2\,-\,1)}}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2}$, $\{\phi_{i,n_2}p_{n_2+1}\}_{i\,=\,n_2\,-\,n_1\,+\,1,\ldots,\,n_2}$
A,T/A	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,\,n_2\,-\,1\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2\,-\,1)}}$, $\{p_i\}_{i\,=\,2,\ldots,\,n_2}$, $\phi_{n_2,n_2}p_{n_2+1}$
A,T/A,T	$\{\phi_{i,j}\}_{\substack{i\,=\,1,\ldots,\,n_2\,-\,1\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2\,-\,1)}}$, $\{p_{i,j}\}_{\substack{i\,=\,2,\ldots,\,n_2\\ j\,=\,i,\ldots,\,\min(n_1\,+\,i\,-\,1,\,n_2)}}$, $\{\phi_{i,n_2}p_{i+1,n_2+1}\}_{i\,=\,n_2\,-\,n_1\,+\,1,\ldots,\,n_2}$

different occasion of first-capture i where the parameters are also dependent on the recapture occasion of the study the animal was recaptured at. Example 3.7 shows how parameter redundancy can be examined in a capture-recapture model with separate juvenile survival probabilities using the `Maple` procedure `capmodJintrinsic`.

Example 3.7 - The capture-recapture $T^1/C/A,T$ intrinsic model: [See electronic appendix `example3.7.mw`] Consider the T^1/C/A,T model with time-dependent first occasion survival probabilities, constant adult survival as well as age- and time-dependent recapture probabilities, where there are four first-capture occasions and four recapture occasions. Observe that we illustrate this `Maple` code by using an example which assumes the age of the animals is known where this assumption may not be applicable for all capture-recapture studies. We can find obtain a simpler exhaustive summary

for the model by using the following `Maple` code:

```
P := capmodJintrinsic(2,1,1,4,4,4);
# Inputs: (x,J,y,z,n1,n2);
# x = juvenile survival probability; J = number of juvenile occasions;
# y = adult survival probability; z = recapture probability;
# for x, y and z: '1'=C, '2'=T, '3'=A, '4'=A,T: x can only be C or T;
# n1 = number of first-capture occasions;
# n2 = number of recapture occasions.
```

`Maple` returns the matrix

$$\mathbf{P} = \begin{bmatrix} \phi_{1,4}p_{2,5} & \phi_{1,3}p_{2,4} & \phi_{1,2}p_{2,3} & \phi_{1,1}p_{2,2} \\ 0 & \phi_a p_{3,5} & \phi_a p_{3,4} & \phi_a p_{3,3} \\ 0 & 0 & \phi_a p_{4,5} & \phi_a p_{4,4} \\ 0 & 0 & 0 & \phi_a p_{5,5} \\ 0 & \phi_{1,3}(1-p_{2,4}) & \phi_{1,2}(1-p_{2,3}) & \phi_{1,1}(1-p_{2,2}) \\ 0 & 0 & \phi_a(1-p_{3,4}) & \phi_a(1-p_{3,3}) \\ 0 & 0 & 0 & \phi_a(1-p_{4,4}) \end{bmatrix}.$$

We alter the notation to display ϕ_a if the probability of adult survival is constant unlike just ϕ in the previous y/z capture-recapture model. From this matrix $\mathbf{P}$ we can then let $\boldsymbol{\kappa}$ be the vector of all the non-zero terms in matrix $\mathbf{P}$. The parameter deficiency of the model can then be found as described in Section 2.2.

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta):
r := Rank(DD); d := Dimension(theta)-r;
```

The exhaustive summary is

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_{1,4}p_{2,5} \\ \phi_{1,3}p_{2,4} \\ \phi_{1,2}p_{2,3} \\ \phi_{1,1}p_{2,2} \\ \phi_a p_{3,5} \\ \vdots \\ \phi_a(1-p_{3,3}) \\ \phi_a(1-p_{4,4}) \end{bmatrix},$$

with parameters $\boldsymbol{\theta} = [\phi_a,\ \phi_{1,1},\ \phi_{1,2},\ \phi_{1,3},\ \phi_{1,4},\ p_{2,2},\ p_{2,3},\ p_{2,4},\ p_{2,5},\ p_{3,3},\ p_{3,4},\ p_{3,5},\ p_{4,4},\ p_{4,5},\ p_{5,5}]^T$. The derivative matrix is given in the `Maple` file `example3.7.mw` only to save space. This derivative matrix has the model rank and parameter deficiency

$\mathbf{r} := \mathbf{14}, \mathbf{d} := \mathbf{1}$.

The T^1/C/A,T capture-recapture model where there are four first-capture occasions and four recapture occasions has 14 estimable parameter combinations and a parameter deficiency of 1. □

We now provide two tables for capture-recapture models with juvenile survival probabilities: Table 3.5a is for the intrinsic parameter redundancy results when $J = 1$ and Table 3.5b is for the intrinsic parameter redundancy results for a general J. Note that the results from 3.5a can be deduced from Table 3.5b as $B = n_2$ and $G = 0$ when $J = 1$, except for a few cases where the models increase or decrease their model rank and parameter deficiency by 1 if $n_2 > n_1$ when $J = 1$ (these are pointed out in Table 3.5a indicated by † and ‡). In Table 3.5a, the model rank and parameter deficiency are then given in columns two and three respectively, with the final column referring to the method of proof used as detailed in Section 2.3. Table 3.5b gives the model rank and two different parameter deficiencies distinguishing between when $n_2 < n_1 + J$ and when $n_2 \geq n_1 + J$. Most practical studies will have $n_2 < n_1 + J$ as rarely will there be a large difference between the number of first-capture and recapture occasions. There are n_1 first-capture occasions and n_2 recapture occasions, where $n_2 \geq 3$ with $n_2 \geq n_1$. The results of Tables 3.5a and 3.5b assume there are perfect data observed with at least one observation of each possible distinct capture-history.

As before, we give a list of estimable parameter combinations for different $x^J/y/z$ capture-recapture models, as given in Table 3.6. Any model using the reparameterisation theorem proof is based on the reparameterisations given in Table 3.6. The estimable parameter combinations for the $\mathrm{C}^J/\mathrm{A}/z$ models can be found in Table 3.4 as the models are identical to A/z models as both juvenile and adult survival probabilities are then age-dependent. This is also the case for the $\mathrm{T}^J/\mathrm{A,T}/z$ models as they are identical to the $\mathrm{A,T}/z$ models as both juvenile and adult survival probabilities are then age- and time-dependent.

Table 3.5a: Intrinsic parameter redundancies for capture-recapture $x^1/y/z$ models with first-year survival probabilities

Model	Rank	Deficiency	Method
C^1/C/C	3	0	Extension Theorem
C^1/C/T	n_2+2	0	Extension Theorem
C^1/C/A	n_2+2	0	Extension Theorem
C^1/C/A,T	$E+2$	0	Extension Theorem
C^1/T/C	n_2+1	0	Extension Theorem
C^1/T/T †	$2n_2$	0	Extension Theorem (for $d=0$†)
C^1/T/A	$2n_2$	0	Extension Theorem
C^1/T/A,T	$E+n_2-1$	1	Reparameterisation
C^1/A/C	n_2+1	0	Identical to A/C model
C^1/A/T	$2n_2$	0	Identical to A/T model
C^1/A/A	$2n_2-1$	1	Identical to A/A model
C^1/A/A,T	$E+n_2-1$	1	Identical to A/A,T model
C^1/A,T/C	$E-n_1+2$	0	Extension Theorem
C^1/A,T/T †	$E-n_1+n_2+1$	0	Extension Theorem (for $d=0$†)
C^1/A,T/A	$E-n_1+n_2$	1	Reparameterisation
C^1/A,T/A,T †	$2E-2n_1+2$	n_1+1	Reparameterisation
T^1/C/C	n_1+2	0	Extension Theorem
T^1/C/T	n_1+n_2+1	0	Extension Theorem
T^1/C/A	n_1+n_2+1	0	Extension Theorem
T^1/C/A,T ‡	$E+n_1$	1	Reparameterisation (for $d=1$‡)
T^1/T/C	n_1+n_2	0	Extension Theorem
T^1/T/T	n_1+2n_2-2	1	Reparameterisation
T^1/T/A	n_1+2n_2-1	0	Extension Theorem
T^1/T/A,T ‡	$E+n_1+n_2-3$	2	Reparameterisation
T^1/A/C	n_1+n_2	0	Extension Theorem
T^1/A/T	n_1+2n_2-1	0	Extension Theorem
T^1/A/A	n_1+2n_2-2	1	Reparameterisation
T^1/A/A,T ‡	$E+n_1+n_2-3$	2	Reparameterisation
T^1/A,T/C	$E+1$	0	Identical to A,T/C model
T^1/A,T/T	$E+n_2-1$	1	Identical to A,T/T model
T^1/A,T/A	$E+n_2-1$	1	Identical to A,T/A model
T^1/A,T/A,T	$2E-n_1$	n_1	Identical to A,T/A,T model

†: When $n_2 > n_1$ the rank decreases by 1 and the deficiency increases by 1;
‡: When $n_2 > n_1$ the rank increases by 1 and the deficiency decreases by 1;
$E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.

Table 3.5b: Intrinsic parameter redundancies for capture-recapture $x^J/y/z$ models for a general number of juvenile years J

Model	Rank	Deficiency (1)	Deficiency (2)
C^J/C/C	$J+2$	0	0
C^J/C/T	n_2+J+1	0	0
C^J/C/A	n_2+J+1	0	0
C^J/C/A,T	$E+J+1$	0	0
C^J/T/C	n_2+1	0	0
C^J/T/T	$2n_2-1$	0	1
C^J/T/A	$2n_2$	0	0
C^J/T/A,T	$E+n_2-1$	1	1
C^J/A/C	n_2+1	0	0
C^J/A/T	$2n_2$	0	0
C^J/A/A	$2n_2-1$	1	1
C^J/A/A,T	$E+n_2-1$	1	1
C^J/A,T/C	$E-n_1+n_2-B+J+1+G$	0	0
C^J/A,T/T	$E-n_1+2n_2-B+J+G-d$	0	1
C^J/A,T/A	$E-n_1+2n_2-B+J-1+G$	1	1
C^J/A,T/A,T	$2E-n_1+n_2-B+J+G-d$	n_2-J	n_1
T^J/C/C	$B+n_1-n_2+2-G$	0	0
T^J/C/T	$B+n_1+1-G$	0	0
T^J/C/A	$B+n_1+1-G$	0	0
T^J/C/A,T	$E+n_1-n_2+B+1-G-d$	n_1-n_2+J	0
T^J/T/C	$B+n_1-J+1-G$	0	0
T^J/T/T	$B+n_1+n_2-J-1-G$	1	1
T^J/T/A	$B+n_1+n_2-J-G$	0	0
T^J/T/A,T	$E+n_1+B-J-G-d$	n_1-n_2+J+1	1
T^J/A/C	$B+n_1-J+1-G$	0	0
T^J/A/T	$B+n_1+n_2-J-G$	0	0
T^J/A/A	$B+n_1+n_2-J-1-G$	1	1
T^J/A/A,T	$E+n_1+B-J-G-d$	n_1-n_2+J+1	1
T^J/A,T/C	$E+1$	0	0
T^J/A,T/T	$E+n_2-1$	1	1
T^J/A,T/A	$E+n_2-1$	1	1
T^J/A,T/A,T	$2E-n_1$	n_1	n_1

d in the rank column refers to the deficiency given in column 3 or 4;
Deficiency (1) is when $n_2 < n_1 + J$ and Deficiency (2) is when $n_2 \geq n_1 + J$;
$E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$; $B = \frac{1}{2}J(2n_2 - J + 1)$;
$G = \frac{1}{2}[(n_2 - n_1 - 1)^2 + (n_2 - n_1 - 1)]$ when $n_2 < n_1 + J$ and
$G = \frac{1}{2}(J-1)(2n_2 - J - 2n_1)$ when $n_2 \geq n_1 + J$.

Table 3.6: Estimable parameter combinations for capture-recapture $x^J/y/z$ models

Model	Estimable Parameter Combinations
C^J/T/T	$\{\phi_i\}_{i=1,\ldots,n_2-1}$, $\{p_i\}_{i=2,\ldots,n_2}$, $\phi_{n_2}p_{n_2+1}$ (only for $n_2 \geq n_1 + J$)
C^J/T/A,T	$\{\phi_i\}_{i=1,\ldots,n_2-1}$, $\{p_{i,j}\}_{\substack{i=2,\ldots,n_2 \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\{\phi_{n_2}p_{i,n_2+1}\}_{i=\max(n_2-n_1+2,J+2),\ldots,n_2+1}$
C^J/A,T/T	$\{\phi_i\}_{i=1,\ldots,J}$, $\{p_i\}_{i=2,\ldots,n_2}$, $\{\phi_{i,j}\}_{\substack{i=J+1,\ldots,n_2-1 \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, $\{\phi_{i,n_2}p_{n_2+1}\}_{i=n_2-n_1+1,\ldots,J}$ (only for $n_2 \geq n_1 + J$)
C^J/A,T/A	$\{\phi_i\}_{i=1,\ldots,J}$, $\{p_i\}_{i=2,\ldots,n_2}$, $\phi_{n_2,n_2}p_{n_2+1}$, $\{\phi_{i,j}\}_{\substack{i=J+1,\ldots,n_2-1 \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$
C^J/A,T/A,T	$\{\phi_i\}_{i=1,\ldots,J}$, $\{\phi_{i,j}\}_{\substack{i=J+1,\ldots,n_2-1 \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, $\{p_{i,j}\}_{\substack{i=2,\ldots,n_2 \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\{\phi_{i,n_2}p_{i+1,n_2+1}\}_{i=\max(n_2-n_1+1,J+1),\ldots,n_2}$
T^J/C/A,T	$\{\phi_{i,j}\}_{\substack{i=1,\ldots,J \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, ϕ_a, $\{p_{i,j}\}_{\substack{i=2,\ldots,n_2 \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\{\phi_{i,n_2}p_{i+1,n_2+1}\}_{i=\max(n_2-n_1+1,J),\ldots,J}$
T^J/T/T	$\{\phi_{i,j}\}_{\substack{i=1,\ldots,J \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, $\{\phi_i\}_{i=J+1,\ldots,n_2-1}$, $\{p_i\}_{i=2,\ldots,n_2}$, $\phi_{n_2}p_{n_2+1}$; if $n_2 - n_1 < J$ also: $\{\phi_{i,n_2}p_{n_2+1}\}_{i=n_2-n_1+1,\ldots,J}$
T^J/T/A,T	$\{\phi_{i,j}\}_{\substack{i=1,\ldots,J \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, $\{\phi_i\}_{i=J+1,\ldots,n_2-1}$, $\{p_{i,j}\}_{\substack{i=2,\ldots,n_2 \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\{\phi_{n_2}p_{i,n_2+1}\}_{i=J+2,\ldots,n_2+1}$; if $n_2 - n_1 < J$ also: $\{\phi_{i,n_2}p_{i+1,n_2+1}\}_{i=n_2-n_1+1,\ldots,J}$
T^J/A/A	$\{\phi_{i,j}\}_{\substack{i=1,\ldots,J \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\{\phi_i\}_{i=J+1,\ldots,n_2-1}$, $\{p_i\}_{i=2,\ldots,n_2}$, $\phi_{n_2}p_{n_2+1}$
T^J/A/A,T	$\{\phi_{i,j}\}_{\substack{i=1,\ldots,J \\ j=i,\ldots,\min(n_1+i-1,n_2-1)}}$, $\{\phi_i\}_{i=J+1,\ldots,n_2-1}$, $\{p_{i,j}\}_{\substack{i=2,\ldots,n_2 \\ j=i,\ldots,\min(n_1+i-1,n_2)}}$, $\phi_{n_2}p_{n_2+1,n_2+1}$; if $n_2 - n_1 < J$ also: $\{\phi_{i,n_2}p_{i+1,n_2+1}\}_{i=n_2-n_1+1,\ldots,J}$
All other models have no parameter redundancy so all the parameters in those models can be estimated.	

We have now finished our discussion of intrinsic parameter redundancy in capture-recapture models. We now proceed to examine extrinsic parameter redundancy results, illustrated by a data set involving European dippers.

3.7 Extrinsic parameter redundancy: The dippers data set

The issues of having imperfect data, which has been previously highlighted in Section 2.6 for mark-recovery models, can also be an issue in capture-recapture models. The large number of capture-histories in long studies means that unless the number of animals in the study is also large, there are going to be many capture-histories which are not observed during the survey. For example, if the probabilities of survival and recapture are quite small, say $\phi = p = 0.2$, then the probability of the capture-history 11111 is equal to $Pr(11111) = 0.2^{10} = 0.0000001$, so that capture-history would very likely be unobserved in a study. There may also be examples of say 30-year studies where the animal's life expectancy is much less than 30 years, so that the capture-history '111...111' is essentially an impossibility. The procedure given in this section will produce an exhaustive summary from the observed capture-histories where the usual parameter redundancy methods from the first chapter can be used to obtain the rank of the derivative matrix and its parameter deficiency. The full procedure is listed in the `Maple` file `example3.8.mw` and we will present a real data set on European dippers to illustrate its use in practice.

Example 3.8 - Extrinsic parameter redundancy in the capture-recapture model - The European dippers data set: [See electronic appendix `example3.8.mw`] We consider a data set on European dippers in this example. The data were originally published in Marzolin (1988) and then used in a number of practical applications such as in Lebreton et al. (1992), Brooks et al. (2000) and Royle (2008). This data set involved the capture and recapture of European dippers (*Cinclus cinclus*) collected for seven years between 1981 and 1987 by Gilbert Marzolin in eastern France. The data set consists of capturing and recapturing breeding adults each year between its breeding period in early March and the 1st of June. There were a total of 294 animals captured during the study for six recapture occasions, and a representation of the data can be found in Table 3.7 as well as in Lebreton et al. (1992, Table 10).

Table 3.7: Capture-histories from the dippers data set of Marzolin (1988)

Capture-History	Number of males	Number of females	Total number of animals
1111110	1	0	1
1111100	0	1	1
1111000	1	1	2
1101110	0	1	1
1100000	4	2	6
1010000	1	1	2
1000000	5	4	9
0111111	0	2	2
0111110	0	1	1
0111100	1	2	3
0111000	1	1	2
0110110	0	1	1
0110000	7	4	11
0100000	11	18	29
0011111	0	2	2
0011110	1	1	2
0011100	4	2	6
0011000	8	4	12
0010110	1	0	1
0010000	11	18	29
0001111	6	2	8
0001110	3	4	7
0001100	6	5	11
0001011	0	1	1
0001001	1	1	2
0001000	6	10	16
0000111	10	6	16
0000110	3	6	9
0000100	9	7	16
0000011	12	11	23
0000010	11	12	23
Totals:	124	131	255
All the other capture-histories were not observed.			

This data set is quite sparse as there are only 31 distinct capture-histories recorded, compared to the 126 distinct capture-histories there would be in a perfect data set. Furthermore, there are only 24 distinct capture-histories recorded in the males-only data set and only 29 distinct capture-histories recorded in the females-only data set. To examine parameter redundancy in `Maple` we list all these capture-histories in a matrix with each row being a different capture-history and each column representing the capture occasion going from the first-capture occasion in the first column to the

(n_2+1)th capture occasion in the last column (as there is one capture occasion at the beginning of the study where only first capturing takes place and no recaptures are possible). We examine the data set by treating male and female animals separately and we denote **Pmale** to be the matrix of all male capture-histories and **Pfemale** to be the matrix of all female capture-histories. We also examine the combined data set with both male and female capture-histories where we denote **Pall** to be the matrix of all capture-histories regardless of gender. For example, the matrix of all male capture-histories is given by

$$\mathbf{Pmale} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

For the dippers data set **Pmale** is a 24 by 7 matrix, **Pfemale** is a 29 by 7 matrix and **Pall** is a 31 by 7 matrix. We can then put the relevant matrix into the `Maple` procedure supplied in `example3.8.mw` to obtain the final rank and parameter deficiency for this particular data set. As an illustration of how this `Maple` procedure can be used to obtain parameter redundancy results, consider the T/A capture-recapture model for when viewing the males-only data set:

```
kappa := caprecaphistories(Pmale,2,3):
# Inputs: (Data,y,z); Data = Data of all capture-histories;
# y = survival probability; z = recapture probability;
# for y and z: '1'=C, '2'=T, '3'=A, '4'=A,T.
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),theta):
# Note: using the 'logvector' exhaustive summary to speed up computation.
hybrid := Formnum2(DD,theta);
```

This gives the exhaustive summary as

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_1 p_2 \phi_2 p_3 \phi_3 p_4 \phi_4 p_5 \phi_5 p_6 (1 - \phi_6 p_7) \\ \phi_1 p_2 \phi_2 p_3 \phi_3 p_4 (1 - \phi_4 + \phi_4 \bar{p}_5 (1 - \phi_5 + \phi_5 \bar{p}_6 (1 - \phi_6 p_7))) \\ \phi_1 p_2 (1 - \phi_2 + \phi_2 \bar{p}_3 (1 - \phi_3 + \phi_3 \bar{p}_4 (1 - \phi_4 + \phi_4 \bar{p}_5 (1 - \phi_5 + \phi_5 \bar{p}_6 (1 - \phi_6 p_7))))) \\ \vdots \\ 1 - \phi_5 + \phi_5 \bar{p}_2 (1 - \phi_6 p_3) \\ \phi_6 p_2 \\ 1 - \phi_6 p_2 \end{bmatrix},$$

where $\bar{p}_i = 1 - p_i$. The parameters in this model are $\boldsymbol{\theta} = [\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, p_2, p_3, p_4, p_5, p_6, p_7]^T$. The derivative matrix is given in the `Maple` file `example3.8.mw` only to save space. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{12}, \mathbf{d} := \mathbf{0}$.

Therefore, only having these 24 distinct male capture-histories is still sufficient to estimate every parameter in the T/A model for the males-only data. Observe that `Maple` can have memory limitations which are alleviated by the use of logvector form of the exhaustive summary and by using the hybrid symbolic-numerical method instead of the pure symbolic method.

We can now show here the extrinsic parameter redundancy results for all of the 16 different possible y/z capture-recapture models for each of the three data sets of the males-only data, the females-only data, and the combined data. Table 3.8 shows these parameter deficiencies in columns three to five with the model rank in the second column and the final column showing what the intrinsic parameter deficiency would be when we observe perfect data.

We can see from Table 3.8 that for the majority of the simpler models the capture-histories observed allow all the parameters to be estimated. We initially expected that for such a sparse data set with a lower number of distinct capture-histories, more models would be parameter redundant, and the models that were parameter redundant would have had larger deficiencies. It is however hard in most cases to estimate all of the parameters if one set of parameters are age- and time-dependent. □

Table 3.8: Extrinsic parameter redundancies for capture-recapture y/z models for the dippers data set of Marzolin (1988)

Model	Number of Parameters	Parameter deficiency of			
		Males	Females	Combined	Intrinsic
C/C	2	0	0	0	0
C/T	7	0	0	0	0
C/A	7	0	0	0	0
C/A,T	22	2	0	0	0
T/C	7	0	0	0	0
T/T	11	1	1	1	1
T/A	12	0	0	0	0
T/A,T	27	5	2	1	1
A/C	7	0	0	0	0
A/T	12	0	0	0	0
A/A	12	2	1	1	1
A/A,T	27	6	2	2	1
A,T/C	22	2	0	0	0
A,T/T	27	5	2	2	1
A,T/A	27	6	2	2	1
A,T/A,T	42	19	15	13	6

A further point arising from this analysis is whether we can list all the possible capture-histories for certain n_1 and n_2 values as a double-check of our results in Table 3.3. Generating all the possible capture-histories is achievable using the procedure `fullhistcaprecap`, which uses the `combinat` package in `Maple`. This code is shown in Example 3.8 below.

Example 3.8 revisited: [See electronic appendix `example3.8.mw`] Consider a study with $n_1 = 6$ as the number of first-capture occasions and $n_2 = 8$ as the number of recapture occasions, examining the A,T/A,T capture-recapture model. We can obtain a matrix of every possible distinct capture-history from the `Maple` code

```
Data := fullhistcaprecap(6,8):
# Inputs: (n1,n2);
# n1 = number of first-capture occasions;
# n2 = number of recapture occasions.
```

This gives the **Data** matrix containing all the possible capture-histories as

$$\mathbf{Data} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

We can obtain the parameter redundancy of the capture-histories by the usual methods from Section 2.2,

```
kappa := caprecaphistories(Data,4,4):
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

The exhaustive summary, parameter set and derivative matrix are given in the `Maple` file `example3.8.mw` only to save space. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{60}, \mathbf{d} := \mathbf{6}$.

This agrees with our previous results from Table 3.3 as the rank should be equal to $r = 2E - n_1 = 2 \times 33 - 6 = 60$ and the parameter deficiency should be equal to $d = n_1 = 6$. □

To end this chapter on parameter redundancy in capture-recapture models, we generalise these extrinsic parameter redundancy results in the next section.

3.8 An extrinsic sparseness statistic for capture-recapture models

In this section we identify the possible capture-histories that a data set is likely to have. This approach has been considered in the mark-recovery model in Cole et al. (2012, Table 4) where they consider the extrinsic parameter deficiency of a model given that they have a certain number of m diagonal values from the m-array that have been observed. Consider a capture-recapture model with a medium probability of annual survival with $\phi = 0.5$ and a low probability of recapture with $p = 0.2$ for all parameters. We show the probability of each possible capture-history of an animal that was captured at the first-capture occasion in the second column of Table 3.9. We then show how many animals we expect to observe with that capture-history if we mark $n = 10$, $n = 100$ and $n = 1000$ animals at the first-capture occasion.

Table 3.9: Expected number of capture-histories for medium survival probabilities and low recapture probabilities

	Probability	Expected number of animals recaptured if		
Capture-history	of capture-history	$n = 10$	$n = 100$	$n = 1000$
1000	0.844	8	84	844
1100	0.086	0	8	86
1010	0.036	0	3	36
1001	0.016	0	1	16
1110	0.009	0	0	9
1101	0.004	0	0	4
1011	0.004	0	0	4
1111	0.001	0	0	1

If 100 animals were captured at the first-capture occasion and followed for three recapture occasions, the capture-histories which we expect to see at least once are 1000, 1100, 1010 and 1000. It is seen from these capture-histories that we only expect to see an animal at most on two capture occasions. If say 1000 animals were captured in the first-capture occasions instead, we would then expect at least one occasion of all eight distinct capture-histories to be observed during a study. While in practice not all of the distinct capture-histories may actually be observed, we expect the data to be typically close to the expected number of animals recaptured given in Table 3.9. Note that we use the capture-histories we expect to see at least *once* for our analyses, though a similar analysis could be done for example when we expect to see that

capture-history at least twice.

Let us define a statistic, c, to indicate the difference between the time of the first capture occasion and the last recapture occasion the animal was seen where $0 \leq c < n_2$. Differences in the expected numbers of capture-histories observed varies with the capture-recapture model specified and with the values of ϕ and p, but typically sparser data sets with lower numbers of distinct capture-histories will have lower values of c. Suppose that we have all capture-histories with a difference of c between an animal's first capture and last recapture, we can then calculate the parameter deficiency for only those particular capture-histories. Real data will never have this exact pattern of capture-histories, but we would expect a data set which is very sparse and/or has few recaptures per occasion to behave like a model with a low value of c.

We set $n_1 = n_2$ throughout this section; we have attempted to generalise results to models where $n_2 > n_1$ but this fails to obtain consistent patterns for all of the capture-recapture model possibilities. Table 3.10 shows the parameter deficiency of the model given the number of recapture occasions, n_2, and the maximum difference between the number of occasions between first capture and last recapture, c. The parameter deficiency of the model when $c > 1$ is given in the second column, with the parameter deficiency of the model when $c = 0$ or $c = 1$ in the third column, as well as the intrinsic parameter deficiency in the final column. The results of Table 3.10 were generated by the hybrid symbolic-numerical approach from Choquet and Cole (2012). We also demonstrate how the parameter redundancy results are obtained in Example 3.9 below.

Example 3.9: Consider when the difference between first capture and last recapture is a maximum of two recapture occasions, $c = 2$, for a model where there are four first-capture occasions and four recapture occasions, $n_2 = 4$. This means we only consider the 14 capture-histories {11100, 10100, 11000, 10000, 01110, 01010, 01100, 01000, 00111, 00101, 00110, 00100, 00011, 00010}, and not the 16 capture-histories {11111, 11110, 11101, 11011, 10111, 11010, 10110, 11001, 10101, 10011, 10010, 10001, 01111, 01101, 01011, 01001} as all these capture-histories have either three or four recapture occasions between first capture and last recapture. We can then see whether a certain capture-recapture y/z model has parameter redundancy from Table 3.10. In this case where $c = 2$ and $n_1 = n_2 = 4$, the models C/C, C/T, C/A, T/C, T/A, A/C and A/T are all not parameter redundant. If we observe more capture-histories so that $c = 3$, then the models C/A,T and A,T/C also become full rank. We do not believe that a study will have the exact capture-histories listed where $c = 2$ but a data set will

Table 3.10: Parameter redundancies for capture-recapture y/z models where there is a maximum difference c between the number of occasions between first capture and last recapture

Model	Deficiency when $c > 1$	Deficiency when $c = \{0, 1\}$	Intrinsic
C/C	0	1	0
C/T	0	1	0
C/A	0	1	0
C/A,T	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + 1$	0
T/C	0	1	0
T/T	1	n_2	1
T/A	0	n_2	0
T/A,T	$\frac{1}{2}(n_2 - c)(n_2 - c - 1) + 1$	$\frac{1}{2}(n_2 - c + 2)(n_2 - c + 1)$	1
A/C	0	1	0
A/T	0	n_2	0
A/A	$n_2 - c$	n_2	1
A/A,T	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + 1$	$\frac{1}{2}(n_2 - c + 2)(n_2 - c + 1)$	1
A,T/C	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + 1$	0
A,T/T	$\frac{1}{2}(n_2 - c)(n_2 - c - 1) + 1$	$\frac{1}{2}(n_2 - c + 2)(n_2 - c + 1)$	1
A,T/A	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + 1$	$\frac{1}{2}(n_2 - c + 2)(n_2 - c + 1)$	1
A,T/A,T	$(n_2 - c)^2 + n_2$	n_2^2	n_1

typically have very similar capture-histories to those given here. □

The results of Table 3.10 show that simpler models of C/T, C/A, T/C, T/A, A/C and A/T models have no parameter redundancy even when sparse data sets with low values of c, say when $c = 2$ or 3, are considered. This is an desirable feature to observe as it shows that generally quite sparse data sets have no parameter redundancy.

3.9 Discussion

This chapter began viewing capture-recapture models using m-arrays as a way of representing the data. This however does not allow for age-dependent parameters so we then used individual capture-histories as an alternative way of developing an exhaustive summary for capture-recapture models. The analysis of these capture-histories shows that the majority of capture-recapture models are not intrinsically parameter redundant, which means that inference can be made on all of the parameters in the model. The models that are parameter redundant have estimable parameter combinations that only confound the last set of parameters for the last recapture occasion of the study. The intrinsic parameter redundancy results such as in Tables 3.3, 3.5a and 3.5b

provide the baseline context for the modelling of these ecological studies. However the extrinsic procedures are possibly going to be the most widely used part of this chapter because it is almost impossible to observe every distinct capture-history, especially for a study with a large number of recapture occasions and a low probability of recapture. To show how much the parameter deficiency changes when there is sparse data, we considered a capture-recapture data set on European dippers from Marzolin (1988). Table 3.8 shows extrinsic parameter redundancy results when observing an incomplete data set. It shows that while sparse data sets can still be full rank for some simple capture-recapture models, it can mean that there are large parameter deficiencies in the more complex models, especially when age- and time-dependent parameters are considered. We have also given tables that provide some indication of how sparseness affects parameter deficiency by using a sparseness statistic c in Table 3.10.

The intrinsic parameter redundancy analyses of this chapter are applicable to current research in this area of statistical ecology and our results are published in Hubbard et al. (2014). Some of our results have been identified in previous research, such as in Lebreton et al. (1992, Table 3) where they consider the fully time-dependent Cormack-Jolly-Seber model, and in Catchpole and Morgan (1997, Example 5) where they consider parameter redundancy in a fully age-dependent model. However, a list of various models with different parameter dependencies has never been compiled in a single taxonomy. Pledger et al. (2003, Table 1) do create a list of independent estimable parameters for different Cormack-Jolly-Seber models, but they use simulation as a method of generating results rather than the formal methodology of determining parameter redundancy which is done in Yu et al. (2014). Further work in this area considering parameter redundancy analyses of different capture-recapture models includes stop-over models, which are special cases of Jolly-Seber models from Jolly (1965) and Seber (1965), where these have the advantage of relaxing the condition on marking the animal to begin with so that there are unknown arrival times. These stop-over models are discussed in Matechou et al. (2013) and the parameter redundancy of these models considered in Matechou (2010). There are other capture-recapture models that are yet to considered with regard to parameter redundancy, such as Pollock's robust capture-recapture model to deal with unequal catchability in Pollock (1982) (e.g. see Bailey et al., 2004b, where this model has been used), and capture-recapture models which account for trap effects such as in Pradel and Sanz-Aguilar (2012).

From a purely mathematical point of view, we believe that the proof of Theorem 3.2 in Appendix B.1 gives an attractive exhaustive summary only consisting of the

reparameterised parameters $\mathbf{s}$ and $\mathbf{t}$. In addition the two-stage extension theorem of Theorem 3.1 is a novel addition to the original extension theorem from Catchpole and Morgan (1997). It is flexible enough proof to allow for age- and time-dependence of all parameters in the model. The relevant proofs could give important results for a wide range of models, not just in ecology.

While the Cormack-Jolly-Seber capture-recapture model provides a basis for parameter estimation using live recaptures, it is worth discussing the assumptions made. As in the mark-recovery model in Section 2.1, we list these assumptions here:

1. The animals are of known age 0 when we mark them in the study if an age-dependent model is proposed. A different model for when ages are unknown is in the form of a stop-over model which has been considered in Pledger et al. (2009) and Matechou et al. (2013), and for parameter redundancy analysis in Matechou (2010). Age-dependent mixture models which differentiate between age classes are also considered in McCrea et al. (2013).

2. Technological and/or human errors are considered to be negligible. This includes issues like incorrect individual identification and incorrect data input into computers. Note that this is somewhat different from having missing observations; we assume there are no *incorrect* observations but accommodate for missing observations. If this is not accounted for when incorrect observations are present, it can be a source of parameter bias meaning inaccurate estimates are obtained. Tag loss is a possible source of bias, which has been considered in McDonald et al. (2003), as well as the misidentification of animals considered in Link et al. (2010).

3. First capture does not affect the animal's survival probability. This can appear in capture-recapture studies when animals are either 'trap-happy' or 'try-shy' when they are more or less likely to be recaptured again. The animals also die and are recaptured independently of each other. This is similar to the assumptions from mark-recovery models from Section 2.1.

4. No recovery of dead animals is included in this model.

The last point is one we explore further in the next chapter, where we integrate both mark-recovery and capture-recapture models into one single model by allowing the collection of both live recaptures and dead recoveries at the same time. This is the capture-recapture-recovery model and is explored in Chapter 4.

Chapter 4

Capture-Recapture-Recovery Models

4.1 Capture-recapture-recovery background

In Chapters 2 and 3 we considered parameter redundancy in mark-recovery and capture-recapture models. In this chapter, we explore parameter redundancy in **capture-recapture-recovery models**. These models can be used for studies where data are collected on *both* live recaptures and dead recoveries, rather than just recapture or recovery only data.

The first consideration of a capture-recapture-recovery model is by Mardekian and McDonald (1981), where previously it was common to conduct two separate analyses and compare results, such as in Anderson and Sterling (1974). However, this model was somewhat restrictive as they only consider the *last* recapture point in the study, and Barker (1995) notes that their method is only valid for a restrictive set of assumptions regarding emigration from the study location. At the same time, a model incorporating tag returns of dead birds in the Jolly-Seber model (of Jolly, 1965, and Seber, 1965) was explored in Buckland (1980, 1982). Buckland (1980, 1982) shows that integrating this information on dead recoveries in a single model improves the accuracy of the survival estimates obtained by solely analysing capture-recapture Jolly-Seber experiments.

The capture-recapture-recovery model is sometimes credited to Burnham (1993). In Burnham (1993), he shows how Cormack-Jolly-Seber capture-recapture data and tag-return data can be integrated together to obtain a joint recapture-recovery model. This allows us to use information on both the live recapturing and dead recovery of

animals to improve survival estimate precision, rather than performing two separate analyses for two different data sets. Burnham (1993) further mentions the problem of parameter identifiability in fully time-dependent capture-recapture-recovery models, which is something that is confirmed in our analysis in this chapter. Barker (1997) extends the model of Burnham (1993) to allow the animal to be resighted alive multiple times between live recapture occasions, as well as potentially resighting the animal multiple times before it is recovered dead. The Barker (1997) model is a generalisation of the Burnham (1993) model under the special case of there being temporary animal migration from the study location at random intervals, which also includes information of live animal resighting between recapture occasions. Both Burnham (1993) and Barker (1997) models without dead recoveries are particular cases of the model presented by Jolly (1965). This work was further extended by Catchpole et al. (1998, 2000) to generate a flexible model where age- and/or time-dependent parameters could be considered. Both the Burnham (1993) and Barker (1997) models are special cases of the model shown in Catchpole et al. (1998, 2000). Our capture-recapture-recovery model given later in this chapter is based on this work of Catchpole et al. (1998, 2000).

A number of extensions to the capture-recapture-recovery models presented in Burnham (1993), Barker (1997) and Catchpole et al. (1998, 2000) have been implemented since. This includes the addition location parameters in King and Brooks (2003), and the consideration of different breeding states such as in Kendall et al. (2006), King (2012) and McCrea (2012). Some of the different applications of capture-recapture-recovery models include on gadwalls in Colorado, USA in Szymczak and Rexstad (1991), on herring gulls in Lebreton et al. (1995), and on Soay sheep in Langrock and King (2013). Statistical inference can be executed in the software package `MARK`, see Cooch and White (2014) for more information. Issues such as goodness of fit and model comparison using these capture-recapture-recovery models have also been explored in McCrea et al. (2012, 2013). As we can see by the wide range of literature as well as the recent developments, capture-recapture-recovery models remain an important set of models to use in ecological research.

We begin this chapter by showing a capture-recapture-recovery model in Section 4.2 which uses the probabilities of individual life-histories as the basis for an exhaustive summary. We extend this to obtain intrinsic parameter redundancy by the use of a simpler exhaustive summary in Sections 4.4 and 4.5, as well as exploring some extrinsic parameter redundancy results in Section 4.6. We then display an extension of this model from Burnham (1993, Section 3.2) which estimates the probability of animal

Table 4.1: A summary of the capture-recapture-recovery model probabilities for $n_1 = n_2 = 4$ from McCrea et al. (2013)

Releases	Recapture or recovery occasion								Not seen again
R_1	$m_{1,2}$	$d_{1,1}$	$m_{1,3}$	$d_{1,2}$	$m_{1,4}$	$d_{1,3}$	$m_{1,5}$	$d_{1,4}$	$m_{1,\infty}$
R_2			$m_{2,3}$	$d_{2,2}$	$m_{2,4}$	$d_{2,3}$	$m_{2,5}$	$d_{2,4}$	$m_{2,\infty}$
R_3					$m_{3,4}$	$d_{3,3}$	$m_{3,5}$	$d_{3,4}$	$m_{3,\infty}$
R_4							$m_{4,5}$	$d_{4,4}$	$m_{4,\infty}$

emigration in Section 4.8 and derive parameter redundancy results for that model.

4.2 Life-histories and how they can form an exhaustive summary

Table 4.1 provides one representation of the data that is observed in a capture-recapture-recovery study (McCrea et al., 2013, and Burnham, 1993). Here R_i represents the number of marked animals released at capture occasion i, $m_{i,j}$ represents the number of animals released at capture occasion i that are next recaptured at capture occasion j, $d_{i,j}$ represents the number of animals released at capture occasion i that are recovered dead at recovery occasion j, and $m_{i,\infty}$ represents the number of animals last seen at capture occasion i that were not recaptured alive or recovered dead during the remainder of the study. Table 4.1 shows all of the possible recapture and recovery occasions in a capture-recapture-recovery study with four recapture and recovery occasions.

We continue to let n_1 denote the number of first-capture occasions in the study and n_2 denote the number of recapture/recovery occasions in the study, where $n_1 \leq n_2$ must hold as there at least as many first-capture occasions as there are recapture/recovery occasions. Typically in capture-recapture-recovery studies there are T capture and recapture occasions as well as T recovery occasions with $n_1 = n_2 = T - 1$. We assume that the number of recapturing occasions is the same as the number of years when an animal can be recovered dead. We also assume that an animal cannot both be recaptured alive and recovered dead later in the same time period for this model.

Let us continue the notation from Chapters 2 and 3 to let $\phi_{i,j}$ denote the probability that an animal the jth year in its ith year of life given that the animal has survived up to that point, $p_{i,j}$ denote the probability that an animal of age $i-1$ is recaptured at capture occasion j, and $\lambda_{i,j}$ denote the probability that an animal of age $i-1$ is

recovered dead in the jth year of the study. If we consider age-dependent parameters in this model then the animals are marked at age 0.

As was the case in Section 3.4, we can create an exhaustive summary using the probabilities of all the observed histories in a study. As these histories possibly contain both recaptures and recoveries, we denote these as being **life-histories** rather than simply capture-histories in this chapter. We can represent the life-histories as before with a '1' indicating an occasion when the animal was recaptured alive, but we now let a '2' denote an occasion when the animal was recovered dead. As before, the first '1' is also used to represent the occasion when the animal was first-captured. A '0' in the life-history code indicates that the animal was neither recaptured or recovered at that occasion. Several examples of life-histories and their probabilities are given below in Examples 4.1 to 4.3.

Example 4.1: Consider the life-history 11100102 where an animal was first-captured at the first capture occasion, recaptured alive during the second, third and sixth capture occasions as well as being recovered dead at the eighth recovery occasion. If we assume that all parameters are only time-dependent with no age-dependency, the probability of this life-history is

$$Pr(11100102) = \phi_1 p_2 \phi_2 p_3 \phi_3 (1-p_4) \phi_4 (1-p_5) \phi_5 p_6 \phi_6 (1-p_7)(1-\phi_7)\lambda_7.$$

This is an example where the animal was recaptured alive multiple times before it was recovered dead. □

Example 4.2: Consider the life-history 00110200 where an animal was first-captured at the third capture occasion, recaptured alive during the fourth recapture occasion and was then recovered dead at the fifth recovery occasion. If we assume that all parameters are only time-dependent, the probability of this life-history is

$$Pr(00110200) = \phi_3 p_4 \phi_4 (1-p_5)(1-\phi_5)\lambda_5.$$

We observe that after an animal is recovered dead the life-history will contain zeros for the remaining of the history as it is impossible to be recaptured or recovered again after the animal has been recovered dead. □

Example 4.3: Finally consider the life-history 00010110 where the animal was first-captured at the fourth capture occasion, recaptured alive during the sixth and seventh capture occasions, and was not recovered dead at all. If we assume that all parameters are only time-dependent, the probability of this life-history is

$$Pr(00010110) = \phi_4(1-p_5)\phi_5 p_6 \phi_6 p_7 \chi_7,$$

where $\chi_7 = (1-\phi_7)(1-\lambda_7) + \phi_7(1-p_8)$. This χ_7 term represents the probability of being uncertain about the animal after the seventh capture occasion, as the animal may have died and not been recovered, or could be still alive but was not recaptured alive. □

The probability of a particular life-history can be generalised as given in Hubbard et al. (2014). Let δ_k denotes the life-history of the animal at time k, where $\delta_k = 0$ would denote no capture or recovery during recapture occasion k, $\delta_k = 1$ would denote an animal being recaptured alive during recapture occasion k, and $\delta_k = 2$ would denote an animal being recovered dead during recovery occasion k (which is before recapture occasion k). Suppose an animal was first-captured at time a and was last recaptured alive or recovered dead at time b, then the probability associated with a particular life-history, h_X, is

$$Pr(h_X) = \begin{cases} \prod_{k=a+1}^{b} \phi_{k-a,k-1} \left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\} \chi_{b-a+1,b} & \text{if } \delta_b = 1, \\ \prod_{k=a+1}^{b-1} \phi_{k-a,k-1} \left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\} & \text{if } \delta_b = 2, \\ \qquad \times (1-\phi_{b-a,b-1})\lambda_{b-a,b-1} & \end{cases} \tag{4.1}$$

where

$$\chi_{i,j} = (1-\phi_{i,j})(1-\lambda_{i,j}) + \phi_{i,j}(1-p_{i+1,j+1})\chi_{i+1,j+1}, \tag{4.2}$$

with $\chi_{i,n_2} = 1$ for all i. An overall likelihood can be obtained as the product of the N different capture-histories observed as

$$L = \prod_{X=1}^{N} Pr(h_X). \tag{4.3}$$

A suitable exhaustive summary for the capture-recapture-recovery model can be generated containing all the probabilities of the distant life-histories as separate exhaustive summary terms.

4.3 Links to other capture-recapture-recovery likelihoods

While we use the formulation of the model from Hubbard et al. (2014) to generate exhaustive summary terms, it is worth noting there are alternative forms of the capture-recapture-recovery models in other literature. We begin by showing a capture-recapture-recovery model considered in Catchpole et al. (1998). The model parameters are the same as the ones defined in Section 4.2 for the survival, recapture and recovery parameters, though they view the data in terms of cohort rather than age; this has no change on the model if each animal is considered to be of age 0 when the animal is marked at occasion i. The data matrices in the model are $\mathbf{d}_{i,j}$ as the number of animals first-captured at capture occasion i which were found dead between recapture occasions j and $j+1$, $\mathbf{v}_{i,j}$ as the number of animals first-captured at capture occasion i which were recaptured at recapture occasion j and not seen again (which also includes when the animal was last seen when it was marked when $i = j$), $\mathbf{w}_{i,j}$ as the number of animals first-captured at capture occasion i which were recaptured at recapture occasion $j+1$, and $\mathbf{z}_{i,j}$ as the number of animals first-captured at capture occasion i which were not recaptured at recapture occasion $j+1$ but were either recaptured or recovered later on. If we let the probability of an animal aged i surviving recapture occasion j be

$$\alpha_{i,j} = \begin{cases} 0, & \text{for } j < i, \\ 1, & \text{for } j = i, \\ \prod_{s=i}^{j-1} \phi_{i,s}, & \text{for } i+1 \leq j \leq n_2, \end{cases}$$

then the likelihood in Catchpole et al. (1998) is given by

$$L \propto \prod_{i=1}^{n_1} \left[\prod_{j=i}^{n_2-1} \{\alpha_{i,j}(1-\phi_{i,j})\lambda_{i,j}\}^{\mathbf{d}_{i,j}} \prod_{j=i}^{n_2} \{\alpha_{i,j}\chi_{i,j}\}^{\mathbf{v}_{i,j}} \prod_{j=i}^{n_2-1} p_{i+1,j+1}^{\mathbf{w}_{i,j}} \left(1-p_{i+1,j+1}\right)^{\mathbf{z}_{i,j}} \right], \tag{4.4}$$

where

$$\chi_{i,j} = (1-\phi_{i,j})(1-\lambda_{i,j}) + \phi_{i,j}\left(1-p_{i+1,j+1}\right)\chi_{i+1,j+1},$$

for all $i = 1, \ldots, n_2 - 1$ and $j = i, \ldots, \min(n_1 + i - 1, n_2 - 1)$ with $\chi_{i,n_2} = 1$ for all i.

It was further shown in Catchpole et al. (2000) that the likelihood could be re-expressed as:

$$L \propto \prod_{i=1}^{n_1} \left[\prod_{j=1}^{n_2-1} \phi_{i,j}^{\mathbf{w}_{i,j}+\mathbf{z}_{i,j}} p_{i,j+1}^{\mathbf{w}_{i,j}} \left(1-p_{c,j+1}\right)^{\mathbf{z}_{i,j}} \left\{\left(1-\phi_{i,j}\right)\lambda_{i,j}\right\}^{\mathbf{d}_{i,j}} \prod_{j=1}^{n_2} \chi_{i,j}^{\mathbf{v}_{i,j}} \right]. \tag{4.5}$$

Observe that if dead recoveries are now not considered in the Catchpole et al. (2000) likelihood, this reduces to the same Cormack-Jolly-Seber likelihood as shown in Section 3.1 without cohort-dependence. If we do not consider dead recoveries by setting $\mathbf{d}_{i,j} = \mathbf{0}$ and $\lambda_{i,j} = 0$ for all i and j, we get

$$L \propto \prod_{i=1}^{n_1} \left\{ \prod_{j=1}^{n_2-1} \phi_{i,j}^{\mathbf{w}_{i,j}+\mathbf{z}_{i,j}} p_{i,j+1}^{\mathbf{w}_{i,j}} (1-p_{c,j+1})^{\mathbf{z}_{i,j}} \prod_{j=1}^{n_2} \chi_{i,j}^{\mathbf{v}_{i,j}} \right\},$$

with

$$\chi_{i,j} = (1-\phi_{i,j}) + \phi_{i,j}\,(1-p_{i+1,j+1})\,\chi_{i+1,j+1}.$$

Removing cohort-dependency then reduces the matrices $\mathbf{v}_{i,j}$, $\mathbf{w}_{i,j}$ and $\mathbf{z}_{i,j}$ to be the vectors $\mathbf{v}_j$, $\mathbf{w}_j$ and $\mathbf{z}_j$ and leaves us with the likelihood

$$L \propto \prod_{j=1}^{n_2-1} \phi_j^{\mathbf{w}_j+\mathbf{z}_j} p_{j+1}^{\mathbf{w}_j} (1-p_{c,j+1})^{\mathbf{z}_j} \prod_{j=1}^{n_2} \chi_j^{\mathbf{v}_j},$$

where

$$\chi_j = (1-\phi_j) + \phi_j\,(1-p_{j+1})\,\chi_{j+1}.$$

If we relabel the vectors $\mathbf{w}_j$ and $\mathbf{z}_j$ in terms of the vectors $\mathbf{a}_j$ and $\mathbf{c}_j$ as in the Cormack-Jolly-Seber model, and by noting that

$$\begin{aligned}
\frac{1-\chi_j}{\phi_j} &= 1-\chi_{j+1}+\chi_{j+1}p_{j+1} \\
\frac{1-\chi_j}{\phi_j} &= 1-(1-p_{j+1})\,\chi_{j+1} \\
1-\chi_j &= \phi_j-\phi_j\,(1-p_{j+1})\,\chi_{j+1} \\
\chi_j &= (1-\phi_j)-\phi_j\,(1-p_{j+1})\,\chi_{j+1},
\end{aligned}$$

we achieve the same time-dependent likelihood as before in Equation (3.2) where

$$L = \prod_{j=1}^{n_2-1} \phi_j^{\mathbf{v}_j} p_{j+1}^{\mathbf{a}_{j+1}} \chi_j^{\mathbf{c}_j} (1-p_{j+1})^{(\mathbf{v}_j-\mathbf{a}_{j+1})},$$

and

$$\frac{1-\chi_j}{\phi_j} = 1-\chi_{j+1}+\chi_{j+1}p_{j+1},$$

with $\chi_{n_2} = 1$. We can now see there are links between the Catchpole et al. (1998, 2000) likelihoods and the Cormack-Jolly-Seber capture-recapture likelihood of Cormack (1964), Jolly (1965) and Seber (1965).

There are alternative approaches to the creation of a capture-recapture-recovery model such as using a sufficient statistics approach to build up a likelihood for the model based on the work of King and Brooks (2003). We change the notations given in King and Brooks (2003) as they consider differences in location as well as time- and age-dependency in their model. They also further consider cohort-dependence but we will not consider this in our form of the likelihood. The likelihood in King and Brooks (2003) is complex so we will define each separate part of the likelihood to begin with and then give the general form for the model.

The first term in the King and Brooks likelihood given in King and Brooks (2003, Lemma 1) refers to the χ term, which is equal to

$$\chi_{(i,j,k)} = \begin{cases} 1 & (j = k), \\ 1 - \phi_{i,j}\left[1 - (1 - p_{i+1,j+1})\,\chi_{(i+1,j+1,k)}\right] - (1 - \phi_{i,j})\,\lambda_{i,j} & (j < k), \end{cases}$$

where the animal was last seen at age j at time i and it will be age k if the animal survives the end of the study. We have edited the formulae in King and Brooks (2003) to eliminate migration effects by making $\psi(r, s) = 0$ for $r \neq s$ and $\psi(r, r) = 1$ by not having to consider summing over location r or s. We also would like the parameters to be time-dependent, so the subscript i has been inserted into this form of the likelihood. This can be seen to be the same as the Catchpole et al. (1998, 2000) χ term from earlier as

$$\begin{aligned} \chi_{(i,j,k)} &= 1 - \phi_{i,j}\left[1 - (1 - p_{i+1,j+1})\,\chi_{(i+1,j+1,k)}\right] - (1 - \phi_{i,j})\,\lambda_{i,j} \quad (\text{for } j < k) \\ &= 1 - \left[\phi_{i,j} - \phi_{i,j}\,(1 - p_{i+1,j+1})\,\chi_{(i+1,j+1,k)}\right] - (1 - \phi_{i,j})\,\lambda_{i,j} \\ &= (1 - \phi_{i,j}) + \phi_{i,j}\,(1 - p_{i+1,j+1})\,\chi_{(i+1,j+1,k)} - (1 - \phi_{i,j})\,\lambda_{i,j} \\ &= (1 - \phi_{i,j})\,(1 - \lambda_{i,j}) + \phi_{i,j}\,(1 - p_{i+1,j+1})\,\chi_{(i,j+1,k)}. \end{aligned}$$

Let $\mathbf{Q}_{(i,k,j)}$ given as in King and Brooks (2003, Lemma 2) be

$$\mathbf{Q}_{(i,k,j)} = \begin{cases} \phi_{i,k} & (j = k), \\ \phi_{i,j}\,(1 - p_{i+1,j+1})\,\mathbf{Q}_{(i+1,k,j+1)} & (j < k), \end{cases}$$

where an animal at time i of age j is observed until age k. The element of the likelihood which represents the probability of the animal being recaptured alive is given by

$$\mathbf{O}_{(i,k,j)} = p_{i+1,j+1}\mathbf{Q}_{(i,k,j)}.$$

The element of the likelihood which represents the probability of the animal being recovered dead is given in King and Brooks (2003, Lemma 3) to be

$$\mathbf{D}_{(i,k,j)} = \begin{cases} (1-\phi_{i,k})\,\lambda_{i,k} & (j=k), \\ (1-\phi_{i,j})\,\lambda_{i,j}\,(1-p_{i,j})\,\mathbf{Q}_{(i-1,k,j-1)} & (j<k). \end{cases}$$

If we let data matrices in the model be $\mathbf{v}_{(i,k,j)}$ to denote the number of animals that are last recaptured aged j at time i that will be aged $k \geq j$ at the end of the study; $\mathbf{n}_{(i,k,j)}$ to denote the number of animals that are last seen aged k are seen again at age $j+1$ at time i; and $\mathbf{d}_{(i,k,j)}$ to denote the number of animals that are recovered dead between age j and age $j+1$ between time i and $i+1$ that were last observed alive at age $k \leq j$; then the King and Brooks (2003) likelihood is equal to

$$\begin{aligned} L \;=\; \prod_{i=1}^{n_2} \Bigg[& \prod_{j=0}^{n_2} \prod_{k=j}^{n_2} \{\chi_{(i,j,k)}\}^{\mathbf{v}_{(i,j,k)}} \prod_{k=0}^{n_2-1} \prod_{j=k}^{n_2-1} \{\mathbf{D}_{(i,k,j)}\}^{\mathbf{d}_{(i,k,j)}} \\ & \times \prod_{k=0}^{n_2-1} \prod_{j=k}^{n_2-1} \{\mathbf{O}_{(i,k,j)}\}^{\mathbf{n}_{(i,k,j)}} \Bigg]. \end{aligned} \tag{4.6}$$

A further analysis of how the Catchpole et al. (1998, 2000) sufficient data matrices compare with closed-location King and Brooks (2003) sufficient data matrices can also be found in McCrea et al. (2010).

4.4 A simpler exhaustive summary for capture-recapture-recovery models

We have previously given the general probability of a life-history in Equation (4.1). However with this capture-recapture-recovery model, we again have a large number of possible life-histories that can now be observed as shown in Table 4.2. This problem is analogous to the issues we faced in Section 3.4 for the capture-recapture model. What this does is present computational problems in finding the intrinsic parameter redundancy of this model due to the large number of exhaustive summary terms. We deal with this in a similar way to Section 3.4 by deriving an exhaustive summary which is simpler than the exhaustive summary of all the life-histories as their probabilities. This is given in Theorem 4.1 below.

Table 4.2: All the possible life-histories in the capture-recapture-recovery model for $n_1 = n_2 \leq 4$

$n_1 = n_2 = 1$	$n_1 = n_2 = 2$	$n_1 = n_2 = 3$	$n_1 = n_2 = 4$
11 12	011 120	0011 1200	00011 00010 12000 01200
10	111 112	0111 1120	00111 00110 11200 01120
	101 102	0101 1020	00101 01110 10200 01020
	010 012	1111 1112	01111 01010 11120 01112
	110	1011 1102	01011 00100 11020 01102
	100	1101 1012	01101 01100 10120 01012
		1001 1002	01001 01000 10020 01002
		0010 0120	11111 11110 11112 00120
		0110 0112	10111 10110 11102 00112
		0100 0102	11011 11010 11012 00102
		1110 0012	11101 11100 10112 00012
		1010	10011 10010 11002
		1100	10101 10100 10102
		1000	11001 11000 10012
			10001 10000 10002

Theorem 4.1. *A simpler exhaustive summary for the capture-recapture-recovery model consists of the terms:*

- $s_{i,j} = \phi_{i,j} p_{i+1,j+1}$ *for all* $i = 1, \ldots, n_2$ *and* $j = i, \ldots, \min(n_1 + i - 1, n_2)$,
- $t_{i,j} = \phi_{i,j}(1 - p_{i+1,j+1})$ *for all* $i = 1, \ldots, n_2 - 1$ *and* $j = i, \ldots, \min(n_1 + i - 1, n_2 - 1)$,
- *and* $r_{i,j} = (1 - \phi_{i,j})\lambda_{i,j}$ *for all* $i = 1, \ldots, n_2$ *and* $j = i, \ldots, \min(n_1 + i - 1, n_2)$.

The proof of Theorem 4.1 can be found in Appendix B.2 as well as in the supplementary material of Hubbard et al. (2014). The vector of all the **s**, **t** and **r** terms generates a simpler exhaustive summary to obtain the status of intrinsic parameter redundancy in capture-recapture-recovery models. The `Maple` procedure `caprecovmodintrinsic` has been written to generate all the **s**, **t** and **r** terms. This is demonstrated in Example 4.4 below.

We extend previous notation used in Chapters 2 and 3 to capture-recapture-recovery models with $y/(z_1; z_2)$ where y denotes the survival probability, z_1 denotes the recapture probability and z_2 denotes the recovery probability. As previously each probability can be either constant (C), time-dependent (T), age-dependent (A), or

age- and time-dependent (A,T). Therefore, there are 64 different possible capture-recapture-recovery models with different parameter dependencies. As an example, we explore the A/(A,T;T) model below.

Example 4.4 - The intrinsic capture-recapture-recovery A/(A,T;T) model: [See electronic appendix `example4.4.mw`] We can generate a list of simpler exhaustive summary terms and obtain the intrinsic parameter redundancy results for the A/(A,T;T) model, where there are age-dependent survival parameters, age- and time-dependent recapture parameters, and time-dependent recovery parameters. This study has four occasions of capture and recapture/recovery and the model's intrinsic parameter redundancy can be obtained using the `Maple` code below. Observe that we illustrate this `Maple` code by using an example which assumes animals are of known age when marked, however this assumption may not be applicable for all capture-recapture-recovery studies.

```
P := caprecovmodintrinsic(3,4,2,4,4);
# Inputs: (y,z1,z2,n1,n2);
# y = survival probability; z1 = recapture probability;
# z2 = recovery probability;
# for y, z1 and z2: '1'=C, '2'=T, '3'=A, '4'=A,T;
# n1 = number of first-capture occasions;
# n2 = number of recapture/recovery occasions.
```

`Maple` returns the matrix as given below

$$\mathbf{P} = \begin{bmatrix} s_{1,4} & s_{1,3} & s_{1,2} & s_{1,1} \\ 0 & s_{2,4} & s_{2,3} & s_{2,2} \\ 0 & 0 & s_{3,4} & s_{3,3} \\ 0 & 0 & 0 & s_{4,4} \\ 0 & t_{1,3} & t_{1,2} & t_{1,1} \\ 0 & 0 & t_{2,3} & t_{2,2} \\ 0 & 0 & 0 & t_{3,3} \\ r_{1,4} & r_{1,3} & r_{1,2} & r_{1,1} \\ 0 & r_{2,4} & r_{2,3} & r_{2,2} \\ 0 & 0 & r_{3,4} & r_{3,3} \\ 0 & 0 & 0 & r_{4,4} \end{bmatrix} = \begin{bmatrix} \phi_1 p_{2,5} & \phi_1 p_{2,4} & \phi_1 p_{2,3} & \phi_1 p_{2,2} \\ 0 & \phi_2 p_{3,5} & \phi_2 p_{3,4} & \phi_2 p_{3,3} \\ 0 & 0 & \phi_3 p_{4,5} & \phi_3 p_{4,4} \\ 0 & 0 & 0 & \phi_4 p_{5,5} \\ 0 & \phi_1(1-p_{2,4}) & \phi_1(1-p_{2,3}) & \phi_1(1-p_{2,2}) \\ 0 & 0 & \phi_2(1-p_{3,4}) & \phi_2(1-p_{3,3}) \\ 0 & 0 & 0 & \phi_3(1-p_{4,4}) \\ (1-\phi_1)\lambda_4 & (1-\phi_1)\lambda_3 & (1-\phi_1)\lambda_2 & (1-\phi_1)\lambda_1 \\ 0 & (1-\phi_2)\lambda_4 & (1-\phi_2)\lambda_3 & (1-\phi_2)\lambda_2 \\ 0 & 0 & (1-\phi_3)\lambda_4 & (1-\phi_3)\lambda_3 \\ 0 & 0 & 0 & (1-\phi_4)\lambda_4 \end{bmatrix}.$$

The **s**, **t** and **r** terms are represented in the matrix form above for convenience in **P**. The parameter deficiency of the model can be found as described in Section 2.2.

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta):
r := Rank(DD); d := Dimension(theta)-r;
```

If we denote $\boldsymbol{\kappa}$ as the exhaustive summary obtained from all the non-zero terms in the matrix $\mathbf{P}$, this is given as

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_1 p_{2,5} \\ \phi_1 p_{2,4} \\ \phi_1 p_{2,3} \\ \phi_1 p_{2,2} \\ \phi_2 p_{3,5} \\ \vdots \\ (1-\phi_3)\lambda_4 \\ (1-\phi_3)\lambda_3 \\ (1-\phi_4)\lambda_4 \end{bmatrix},$$

with the model's parameters as $\boldsymbol{\theta} = [\phi_1,\ \phi_2,\ \phi_3,\ \phi_4,\ p_{2,2},\ p_{2,3},\ p_{2,4},\ p_{2,5},\ p_{3,3},\ p_{3,4},\ p_{3,5},\ p_{4,4},\ p_{4,5},\ p_{5,5},\ \lambda_1,\ \lambda_2,\ \lambda_3,\ \lambda_4]^T$. The derivative matrix is given in the `Maple` file `example4.4.mw` only to save space. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{18}, \mathbf{d} := \mathbf{0}$.

This means the A/(A,T;T) intrinsic capture-recapture-recovery model with four first-capture occasions and four recapture/recovery occasions is full rank and all of its parameters can in theory be estimated. The extension theorem of Theorem 2.2 can then be used to show that the A/(A,T;T) model is not parameter redundant for any values of $n_1, n_2 \geq 2$. □

A similar procedure, `caprecovmodJintrinsic`, has been written where juvenile survival probabilities are considered along with separate adult survival probabilities. This follows on from work done on separate juvenile survival probabilities in mark-recovery models in Section 2.10 and in capture-recapture models in Section 3.6. In these models the juvenile animals are assumed to have separate survival probabilities compared to the survival probabilities for the adult animals in the study so are modelled using different parameters. We denote these capture-recapture-recovery models

with juvenile survival probabilities, where the animal is a juvenile for J recapture occasions (for $1 \leq J < n_2 - 1$) as $x^J/y/(z_1; z_2)$, where x denotes whether the juvenile survival probabilities are constant or time-dependent for J recapture occasions, y denotes the adult survival probabilities, z_1 denotes the recapture probabilities, and z_2 denotes the recovery probabilities. If x is constant, the model has the survival parameters $\phi_1, \phi_2, \ldots, \phi_J$ for each of the J different recapture occasions, and if x is time-dependent, then the model has the survival parameters $\phi_{i,1}, \phi_{i,2}, \ldots, \phi_{i,J}$ for each different occasion of first-capture i where the parameters are also dependent on an animal's recapture occasion. This is illustrated in Example 4.5 below.

Example 4.5 - The capture-recapture-recovery $T^1/A/(A,T;C)$ intrinsic model: [See electronic appendix `example4.5.mw`] We can generate a list of simpler exhaustive summary terms and obtain the intrinsic parameter redundancy results for the $\mathrm{T}^1/\mathrm{A}/(\mathrm{A,T;C})$ model, where there are time-dependent first-year survival parameters, age-dependent adult survival parameters, age- and time-dependent recapture parameters, and constant recovery parameters. This study has four occasions of first-capture and recapture/recovery and the model's intrinsic parameter redundancy can be obtained using the `Maple` code below. Observe that we illustrate this `Maple` code by using an example which assumes age-dependence is valid whereas this assumption may not be realistically applicable for all capture-recapture-recovery studies, particularly for adult age-dependent survival parameters.

```
P := caprecovmodJintrinsic(2,1,3,4,1,4,4);
# Inputs: (x,J,y,z1,z2,n1,n2);
# x = juvenile survival probability; J = number of juvenile occasions;
# y = adult survival probability; z1 = recapture probability;
# z2 = recovery probability;
# for x, y, z1 and z2: '1'=C, '2'=T, '3'=A, '4'=A,T
#  (x can only be C or T); n1 = number of first-capture occasions;
# n2 = number of recapture/recovery occasions.
```

`Maple` returns the matrix as given below

$$\mathbf{P} = \begin{bmatrix} s_{1,4} & s_{1,3} & s_{1,2} & s_{1,1} \\ 0 & s_2 & s_2 & s_2 \\ 0 & 0 & s_3 & s_3 \\ 0 & 0 & 0 & s_4 \\ 0 & t_{1,3} & t_{1,2} & t_{1,1} \\ 0 & 0 & t_2 & t_2 \\ 0 & 0 & 0 & t_3 \\ r_{1,4} & r_{1,3} & r_{1,2} & r_{1,1} \\ 0 & r_2 & r_2 & r_2 \\ 0 & 0 & r_3 & r_3 \\ 0 & 0 & 0 & r_4 \end{bmatrix} = \begin{bmatrix} \phi_{1,4}p_{2,5} & \phi_{1,3}p_{2,4} & \phi_{1,2}p_{2,3} & \phi_{1,1}p_{2,2} \\ 0 & \phi_2 p_{3,5} & \phi_2 p_{3,4} & \phi_2 p_{3,3} \\ 0 & 0 & \phi_3 p_{4,5} & \phi_3 p_{4,4} \\ 0 & 0 & 0 & \phi_4 p_{5,5} \\ 0 & \phi_{1,3}(1-p_{2,4}) & \phi_{1,2}(1-p_{2,3}) & \phi_{1,1}(1-p_{2,2}) \\ 0 & 0 & \phi_2(1-p_{3,4}) & \phi_2(1-p_{3,3}) \\ 0 & 0 & 0 & \phi_3(1-p_{4,4}) \\ (1-\phi_{1,4})\lambda & (1-\phi_{1,3})\lambda & (1-\phi_{1,2)}\lambda & (1-\phi_{1,1})\lambda \\ 0 & (1-\phi_2)\lambda & (1-\phi_2)\lambda & (1-\phi_2)\lambda \\ 0 & 0 & (1-\phi_3)\lambda & (1-\phi_3)\lambda \\ 0 & 0 & 0 & (1-\phi_4)\lambda \end{bmatrix}.$$

The **s**, **t** and **r** terms are represented in the matrix form above for convenience in **P**. The parameter deficiency of the model can be found as described in Section 2.2.

```
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(kappa,theta):
r := Rank(DD); d := Dimension(theta)-r;
```

If we denote $\boldsymbol{\kappa}$ as the exhaustive summary obtained from all the non-zero terms in the matrix **P**, this is given as

$$\boldsymbol{\kappa} = \begin{bmatrix} \phi_{1,4}p_{2,5} \\ \phi_{1,3}p_{2,4} \\ \phi_{1,2}p_{2,3} \\ \phi_{1,1}p_{2,2} \\ \phi_2 p_{3,5} \\ \vdots \\ (1-\phi_3)\lambda \\ (1-\phi_3)\lambda \\ (1-\phi_4)\lambda \end{bmatrix},$$

with the model's parameters as $\boldsymbol{\theta} = [\phi_{1,1},\ \phi_{1,2},\ \phi_{1,3},\ \phi_{1,4},\ \phi_2,\ \phi_3,\ \phi_4,\ p_{2,2},\ p_{2,3},\ p_{2,4},\ p_{2,5},\ p_{3,3},\ p_{3,4},\ p_{3,5},\ p_{4,4},\ p_{4,5},\ p_{5,5},\ \lambda]^T$. The derivative matrix is given in the `Maple` file `example4.5.mw` only to save space. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{18}, \mathbf{d} := \mathbf{0}$.

The T^1/A/(A,T;C) intrinsic capture-recapture-recovery model with four first-capture occasions and four recapture/recovery occasions is not parameter redundant and all its parameters in theory can be estimated. The extension theorem of Theorem 2.2 can then be used to show that the T^1/A/(A,T;C) model is not parameter redundant for any values of $n_1, n_2 \geq 3$. □

4.5 Intrinsic parameter redundancy results

In this section we present the intrinsic parameter redundancy results for the wide range of capture-recapture-recovery models presented in the previous section. These results are obtained using Theorem 3.2 to generate a simpler exhaustive summary and then `Maple` code to find the rank of a model's derivative matrix. A further result relating the parameter redundancy results for capture-recapture models from Section 3.5 to the results given here for capture-recapture-recovery models, is given in Theorem 4.2 below which shows a simpler method of proof for some of our results.

Theorem 4.2. *If the capture-recapture y/z_1 model is full rank, then the capture-recapture-recovery $y/(z_1; z_2)$ model with the same y and z_1, for any z_2, is also full rank.*

Proof. Consider the exhaustive summary for the capture-recapture-recovery $y/(z_1; z_2)$ model as consisting of two parts. The first part, $\boldsymbol{\kappa}_1$ consists of the terms $s_{i,j} = \phi_{i,j}p_{i+1,j+1}$ for all $i = 1, \ldots, n_2$ and $j = i, \ldots, \min(n_1 + i - 1, n_2)$ and $t_{i,j} = \phi_{i,j}(1 - p_{i+1,j+1})$ for all $i = 1, \ldots, n_2 - 1$ and $j = i, \ldots, \min(n_1 + i - 1, n_2 - 1)$, such that

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} s_{1,1} \\ s_{1,2} \\ s_{2,2} \\ s_{1,3} \\ \vdots \\ t_{1,1} \\ t_{1,2} \\ t_{2,2} \\ t_{1,3} \\ \vdots \end{bmatrix} = \begin{bmatrix} \phi_{1,1}p_{2,2} \\ \phi_{1,2}p_{2,3} \\ \phi_{2,2}p_{3,3} \\ \phi_{1,3}p_{2,4} \\ \vdots \\ \phi_{1,1}(1 - p_{2,2}) \\ \phi_{1,2}(1 - p_{2,3}) \\ \phi_{2,2}(1 - p_{3,3}) \\ \phi_{1,3}(1 - p_{2,4}) \\ \vdots \end{bmatrix}.$$

The second part, $\boldsymbol{\kappa}_2$, consists of the terms $r_{i,j} = (1-\phi_{i,j})\lambda_{i,j}$ for all $i = 1, \ldots, n_2$ and $j = i, \ldots, \min(n_1 + i - 1, n_2)$, such that

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} r_{1,1} \\ r_{1,2} \\ r_{2,2} \\ r_{1,3} \\ \vdots \end{bmatrix} = \begin{bmatrix} (1-\phi_{1,1})\lambda_{1,1} \\ (1-\phi_{1,2})\lambda_{1,2} \\ (1-\phi_{2,2})\lambda_{2,2} \\ (1-\phi_{1,3})\lambda_{1,3} \\ \vdots \end{bmatrix}.$$

Let the parameter vector $\boldsymbol{\theta}_1$ consist of all the parameters $\phi_{i,j}$ and $p_{i,j}$, and the parameter vector $\boldsymbol{\theta}_2$ consist of all the parameters $\lambda_{i,j}$. As the capture-recapture y/z_1 model is full rank then $\mathbf{D}_1 = [\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_1]$ is full rank. The derivative matrix $\mathbf{D}_2 = [\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$ is equal to

$$\mathbf{D}_2 = \left[\frac{\partial\boldsymbol{\kappa}_2}{\partial\boldsymbol{\theta}_2}\right] = \begin{bmatrix} -\phi_{1,1} & 0 & 0 & 0 & \cdots \\ 0 & -\phi_{1,2} & 0 & 0 & \cdots \\ 0 & 0 & -\phi_{2,2} & 0 & \cdots \\ 0 & 0 & 0 & -\phi_{1,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

which consists of the terms $-\phi_{i,j}$ on the diagonal and 0 elsewhere. As long as all $\phi_{i,j}$ are non-zero this will always be full rank. Then as $\mathbf{D}_1$ and $\mathbf{D}_2$ are full rank, by the extension theorem the capture-recapture-recovery $y/(z_1; z_2)$ model is then full rank. □

Theorem 4.2 can also be used to prove results about capture-recapture-recovery models involving juvenile survival probabilities. It can be seen that if the capture-recapture $x^J/y/z_1$ capture-recapture model is full rank then the $x^J/y/(z_1; z_2)$ capture-recapture-recovery model is also full rank. This proof of this is extremely similar to the proof given above where $\boldsymbol{\kappa}_1$ is full rank and $\boldsymbol{\kappa}_2$ is also be full rank as only the parameters $\phi_{i,j}$ change but the derivative matrix will still consist of the terms $-\phi_{i,j}$ on the diagonal and 0 elsewhere.

The general intrinsic parameter redundancy results are given in Tables 4.3a and 4.3b for different $y/(z_1; z_2)$ capture-recapture-recovery models where the second column refers to the rank of the model, which is the number of estimable parameters, and the third column refers to the parameter deficiency of the model where the model is parameter redundant when $d > 0$. There are n_1 first-capture occasions and n_2 re-

capture/recovery occasions, where $n_2 \geq 2$ with $n_2 \geq n_1$. The final column refers to the method of proof used as described in Section 2.3 or when the 'full rank theorem' of Theorem 4.2 is used. The results of Tables 4.3a and 4.3b assume that perfect data are observed with at least one observation of each possible distinct life-history.

For parameter redundant models, we report which parameters are confounded and not estimable in Table 4.4. This table only shows the confounded parameter combinations so that all of the other parameters in the model are estimable, e.g. for the T/(T;T) model the parameters $\phi_{n_2} p_{n_2+1}$ and $(1-\phi_{n_2})\lambda_{n_2}$ are confounded, meaning that the parameters ϕ_i, p_{i+1} and λ_i are all in theory estimable but only for $i = 1, \ldots, n_2 - 1$. Any model which has results which have been proved using the reparameterisation theorem is based on the confounded parameters given in Table 4.4.

Tables 4.3a and 4.3b show that a large proportion of the capture-recapture-recovery models are not parameter redundant and all their parameters are estimable in theory given perfect data. Table 4.4 further shows that for the models which are parameter redundant, only the parameters that refer to the survival probability of the final recapture occasion and the probabilities of recapture and recovery for the final occasion are confounded. This is even true for the complex fully age- and time-dependent case where only the parameters ϕ_{i,n_2}, p_{i+1,n_2+1} and λ_{i,n_2} are confounded for $i = n_2 - n_1 + 1, \ldots, n_2$. Consequently, this means that all the parameters excluding the probabilities of survival until the final occasion's recapture and the recapture and recovery probabilities for the final recapture occasion are in theory estimable for a perfect data set.

The taxonomy of intrinsic parameter redundancy results is extended to capture-recapture-recovery models where we now consider a different set of first-year survival probabilities in Tables 4.5a, 4.5b and 4.5c. We denote these models as $x^1/y/(z_1; z_2)$, where x^1 denotes first occasion survival probabilities as being either constant or time-dependent, y denotes adult survival probability, z_1 denotes recapture probability and z_2 denotes recovery probability. We only consider one occasion of juvenile survival probabilities in these tables though a model with more than one occasion of juvenile survival can be examined using the `Maple` code in the online supplementary file `example4.5.mw`. The final column refers to the method of proof used as described in Section 2.3 though the 'full rank theorem' of Theorem 4.2 can also be used as a method of proof for some of the models as discussed previously. The results of Tables 4.5a, 4.5b and 4.5c assume that perfect data are observed with at least one observation of each possible distinct life-history.

Table 4.3a: Intrinsic parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models (Table 1)

Model	Rank	Deficiency	Method of Proof
C/(C;C)	3	0	Full Rank Theorem
C/(C;T)	$n_2 + 2$	0	Full Rank Theorem
C/(C;A)	$n_2 + 2$	0	Full Rank Theorem
C/(C;A,T)	$E + 2$	0	Full Rank Theorem
C/(T;C)	$n_2 + 2$	0	Full Rank Theorem
C/(T;T)	$2n_2 + 1$	0	Full Rank Theorem
C/(T;A)	$2n_2 + 1$	0	Full Rank Theorem
C/(T;A,T)	$E + n_2 + 1$	0	Full Rank Theorem
C/(A;C)	$n_2 + 2$	0	Full Rank Theorem
C/(A;T)	$2n_2 + 1$	0	Full Rank Theorem
C/(A;A)	$2n_2 + 1$	0	Full Rank Theorem
C/(A;A,T)	$E + n_2 + 1$	0	Full Rank Theorem
C/(A,T;C)	$E + 2$	0	Full Rank Theorem
C/(A,T;T)	$E + n_2 + 1$	0	Full Rank Theorem
C/(A,T;A)	$E + n_2 + 1$	0	Full Rank Theorem
C/(A,T;A,T)	$2E + 1$	0	Full Rank Theorem
T/(C;C)	$n_2 + 2$	0	Full Rank Theorem
T/(C;T)	$2n_2 + 1$	0	Full Rank Theorem
T/(C;A)	$2n_2 + 1$	0	Full Rank Theorem
T/(C;A,T)	$E + n_2 + 1$	0	Full Rank Theorem
T/(T;C)	$2n_2 + 1$	0	Extension Theorem
T/(T;T)	$3n_2 - 1$	1	Reparameterisation
T/(T;A)	$3n_2$	0	Extension Theorem
T/(T;A,T)	$E + 2n_2 - 1$	1	Reparameterisation
T/(A;C)	$2n_2 + 1$	0	Full Rank Theorem
T/(A;T)	$3n_2$	0	Full Rank Theorem
T/(A;A)	$3n_2$	0	Full Rank Theorem
T/(A;A,T)	$E + 2n_2$	0	Full Rank Theorem
T/(A,T;C)	$E + n_2 + 1$	0	Extension Theorem
T/(A,T;T)	$E + 2n_2 - 1$	1	Reparameterisation
T/(A,T;A)	$E + 2n_2$	0	Extension Theorem
T/(A,T;A,T)	$2E + n_2 - 1$	1	Reparameterisation
$E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.			

Table 4.3b: Intrinsic parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models (Table 2)

Model	Rank	Deficiency	Method of Proof
A/(C;C)	$n_2 + 2$	0	Full Rank Theorem
A/(C;T)	$2n_2 + 1$	0	Full Rank Theorem
A/(C;A)	$2n_2 + 1$	0	Full Rank Theorem
A/(C;A,T)	$E + n_2 + 1$	0	Full Rank Theorem
A/(T;C)	$2n_2 + 1$	0	Full Rank Theorem
A/(T;T)	$3n_2$	0	Full Rank Theorem
A/(T;A)	$3n_2$	0	Full Rank Theorem
A/(T;A,T)	$E + 2n_2$	0	Full Rank Theorem
A/(A;C)	$2n_2 + 1$	0	Extension Theorem
A/(A;T)	$3n_2$	0	Extension Theorem
A/(A;A)	$3n_2 - 1$	1	Reparameterisation
A/(A;A,T)	$E + 2n_2 - 1$	1	Reparameterisation
A/(A,T;C)	$E + n_2 + 1$	0	Extension Theorem
A/(A,T;T)	$E + 2n_2$	0	Extension Theorem
A/(A,T;A)	$E + 2n_2 - 1$	1	Reparameterisation
A/(A,T;A,T)	$2E + n_2 - 1$	1	Reparameterisation
A,T/(C;C)	$E + 2$	0	Full Rank Theorem
A,T/(C;T)	$E + n_2 + 1$	0	Full Rank Theorem
A,T/(C;A)	$E + n_2 + 1$	0	Full Rank Theorem
A,T/(C;A,T)	$2E + 1$	0	Full Rank Theorem
A,T/(T;C)	$E + n_2 + 1$	0	Extension Theorem
A,T/(T;T)	$E + 2n_2$	0	Extension Theorem
A,T/(T;A)	$E + 2n_2$	0	Extension Theorem
A,T/(T;A,T)	$2E + n_2 - 1$	1	Reparameterisation
A,T/(A;C)	$E + n_2 + 1$	0	Extension Theorem
A,T/(A;T)	$E + 2n_2$	0	Extension Theorem
A,T/(A;A)	$E + 2n_2 - 1$	1	Reparameterisation
A,T/(A;A,T)	$2E + n_2 - 1$	1	Reparameterisation
A,T/(A,T;C)	$2E + 1$	0	Extension Theorem
A,T/(A,T;T)	$2E + n_2 - 1$	1	Reparameterisation
A,T/(A,T;A)	$2E + n_2 - 1$	1	Reparameterisation
A,T/(A,T;A,T)	$3E - n_1$	n_1	Reparameterisation
$E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.			

Table 4.4: List of confounded parameter combinations for capture-recapture-recovery $y/(z_1; z_2)$ models

Model	Confounded Parameters
T/(T;T)	$\phi_{n_2} p_{n_2+1}$, $(1-\phi_{n_2})\lambda_{n_2}$
T/(T;A,T)	$\phi_{n_2} p_{n_2+1}$, $\{(1-\phi_{n_2})\lambda_{i,n_2}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$
T/(A,T;T)	$\{\phi_{n_2} p_{i+1,n_2+1}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$, $(1-\phi_{n_2})\lambda_{n_2}$
T/(A,T;A,T)	$\{\phi_{n_2} p_{i+1,n_2+1}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$, $\{(1-\phi_{n_2})\lambda_{i,n_2}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$
A/(A;A)	$\phi_{n_2} p_{n_2+1}$, $(1-\phi_{n_2})\lambda_{n_2}$
A/(A;A,T)	$\phi_{n_2} p_{n_2+1}$, $(1-\phi_{n_2})\lambda_{n_2,n_2}$
A/(A,T;A)	$\phi_{n_2} p_{n_2+1,n_2+1}$, $(1-\phi_{n_2})\lambda_{n_2}$
A/(A,T;A,T)	$\phi_{n_2} p_{n_2+1,n_2+1}$, $(1-\phi_{n_2})\lambda_{n_2,n_2}$
A,T/(T;A,T)	$\{\phi_{i,n_2} p_{n_2+1}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$, $\{(1-\phi_{i,n_2})\lambda_{i,n_2}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$
A,T/(A;A)	$\phi_{n_2,n_2} p_{n_2+1}$, $(1-\phi_{n_2,n_2})\lambda_{n_2}$
A,T/(A;A,T)	$\phi_{n_2,n_2} p_{n_2+1}$, $(1-\phi_{n_2,n_2})\lambda_{n_2,n_2}$
A,T/(A,T;T)	$\{\phi_{i,n_2} p_{i+1,n_2+1}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$, $\{(1-\phi_{i,n_2})\lambda_{n_2}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$
A,T/(A,T;A)	$\phi_{n_2,n_2} p_{n_2+1,n_2+1}$, $(1-\phi_{n_2,n_2})\lambda_{n_2}$
A,T/(A,T;A,T)	$\{\phi_{i,n_2} p_{i+1,n_2+1}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$, $\{(1-\phi_{i,n_2})\lambda_{i,n_2}\}_{i\,=\,n_2-n_1+1,\ldots,n_2}$
All other parameters not given in the table are estimable for each model. All models not listed are not parameter redundant and all parameters are estimable.	

4.6 Extrinsic parameter redundancy and the cormorants data set

We obtained intrinsic parameter redundancy results in the previous section and we now extend our analysis to view extrinsic parameter redundancy results in this section. Parameter redundancy can be caused by imperfect data where not all of the distinct life-histories are observed. In real examples, it is unlikely that every possible life-history is observed, and the probability of recording life-histories such as '11111' or '11112' during a study may be extremely small if the survival and recapture/recovery probabilities in the study are small, e.g. if $\phi = 0.2$, $p = 0.2$ and $\lambda = 0.2$ for all survival years and capture/recovery occasions, then the probability of the life-history 11111

Table 4.5a: Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1; z_2)$ models with first-year survival probabilities (Table 1)

Model	Rank	Deficiency	Method of Proof
C^1/C/(C;C)	4	0	Full Rank Theorem
C^1/C/(C;T)	$n_2 + 3$	0	Full Rank Theorem
C^1/C/(C;A)	$n_2 + 3$	0	Full Rank Theorem
C^1/C/(C;A,T)	$E + 3$	0	Full Rank Theorem
C^1/C/(T;C)	$n_2 + 3$	0	Full Rank Theorem
C^1/C/(T;T)	$2n_2 + 2$	0	Full Rank Theorem
C^1/C/(T;A)	$2n_2 + 2$	0	Full Rank Theorem
C^1/C/(T;A,T)	$E + n_2 + 2$	0	Full Rank Theorem
C^1/C/(A;C)	$n_2 + 3$	0	Full Rank Theorem
C^1/C/(A;T)	$2n_2 + 2$	0	Full Rank Theorem
C^1/C/(A;A)	$2n_2 + 2$	0	Full Rank Theorem
C^1/C/(A;A,T)	$E + n_2 + 2$	0	Full Rank Theorem
C^1/C/(A,T;C)	$E + 3$	0	Full Rank Theorem
C^1/C/(A,T;T)	$E + n_2 + 2$	0	Full Rank Theorem
C^1/C/(A,T;A)	$E + n_2 + 2$	0	Full Rank Theorem
C^1/C/(A,T;A,T)	$2E + 2$	0	Full Rank Theorem
C^1/T/(C;C)	$n_2 + 3$	0	Full Rank Theorem
C^1/T/(C;T)	$2n_2 + 2$	0	Full Rank Theorem
C^1/T/(C;A)	$2n_2 + 2$	0	Full Rank Theorem
C^1/T/(C;A,T)	$E + n_2 + 2$	0	Full Rank Theorem
C^1/T/(T;C)	$2n_2 + 1$	0	Extension Theorem
C^1/T/(T;T)†	$3n_2$	0	Extension Theorem ($d = 0$†)
C^1/T/(T;A)	$3n_2$	0	Extension Theorem
C^1/T/(T;A,T)†	$E + 2n_2$	0	Extension Theorem ($d = 0$†)
C^1/T/(A;C)	$2n_2 + 1$	0	Full Rank Theorem
C^1/T/(A;T)	$3n_2$	0	Full Rank Theorem
C^1/T/(A;A)	$3n_2$	0	Full Rank Theorem
C^1/T/(A;A,T)	$E + 2n_2$	0	Full Rank Theorem
C^1/T/(A,T;C)	$E + n_2 + 1$	0	Extension Theorem
C^1/T/(A,T;T)†	$E + 2n_2$	0	Extension Theorem ($d = 0$†)
C^1/T/(A,T;A)	$E + 2n_2$	0	Extension Theorem
C^1/T/(A,T;A,T)	$2E + n_2 - 1$	1	Reparameterisation
C^1/A/(z_1;z_2)	Identical to A/(z_1;z_2) models: See Table 4.3b		
†: when $n_2 > n_1$, rank decreases by 1 and parameter deficiency increases by 1; $E = n_1 n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.			

Table 4.5b: Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1;z_2)$ models with first-year survival probabilities (Table 2)

Model	Rank	Deficiency	Method of Proof
C^1/A,T/(C;C)	$E+2$	0	Full Rank Theorem
C^1/A,T/(C;T)	$E+n_2+1$	0	Full Rank Theorem
C^1/A,T/(C;A)	$E+n_2+1$	0	Full Rank Theorem
C^1/A,T/(C;A,T)	$2E+1$	0	Full Rank Theorem
C^1/A,T/(T;C)	$E+n_2+1$	0	Full Rank Theorem
C^1/A,T/(T;T)	$E+2n_2$	0	Full Rank Theorem
C^1/A,T/(T;A)	$E+2n_2$	0	Full Rank Theorem
C^1/A,T/(T;A,T)	$2E+n_2-1$	0	Full Rank Theorem
C^1/A,T/(A;C)	$E-n_1+n_2+2$	0	Extension Theorem
C^1/A,T/(A;T)	$E-n_1+2n_2+1$	0	Extension Theorem
C^1/A,T/(A;A)	$E-n_1+2n_2$	1	Extension Theorem
C^1/A,T/(A;A,T)	$2E-n_1+n_2$	1	Reparameterisation
C^1/A,T/(A,T;C)	$2E-n_1+1$	0	Extension Theorem
C^1/A,T/(A,T;T)†	$2E-n_1+n_2+1$	0	Extension Theorem ($d=0$†)
C^1/A,T/(A,T;A)	$2E-n_1+n_2$	1	Reparameterisation
C^1/A,T/(A,T;A,T)†	$3E-2n_1+2$	n_1-1	Reparameterisation
T^1/C/(C;C)	n_1+3	0	Full Rank Theorem
T^1/C/(C;T)	n_1+n_2+2	0	Full Rank Theorem
T^1/C/(C;A)	n_1+n_2+2	0	Full Rank Theorem
T^1/C/(C;A,T)	$E+n_1+2$	0	Full Rank Theorem
T^1/C/(T;C)	n_1+n_2+2	0	Full Rank Theorem
T^1/C/(T;T)	n_1+2n_2+1	0	Full Rank Theorem
T^1/C/(T;A)	n_1+2n_2+1	0	Full Rank Theorem
T^1/C/(T;A,T)	$E+n_1+n_2+1$	0	Full Rank Theorem
T^1/C/(A;C)	n_1+n_2+2	0	Full Rank Theorem
T^1/C/(A;T)	n_1+2n_2+1	0	Full Rank Theorem
T^1/C/(A;A)	n_1+2n_2+1	0	Full Rank Theorem
T^1/C/(A;A,T)	$E+n_1+n_2+1$	0	Full Rank Theorem
T^1/C/(A,T;C)	$E+n_1+2$	0	Extension Theorem
T^1/C/(A,T;T)	$E+n_1+n_2+1$	0	Extension Theorem
T^1/C/(A,T;A)	$E+n_1+n_2+1$	0	Extension Theorem
T^1/C/(A,T;A,T)‡	$2E+n_1$	1	Reparameterisation ($d>0$‡)

†: when $n_2 > n_1$, rank decreases by 1 and parameter deficiency increases by 1;
‡: when $n_2 > n_1$, rank increases by 1 and parameter deficiency decreases by 1;
$E = n_1n_2 - \frac{1}{2}n_1^2 + \frac{1}{2}n_1$.

Table 4.5c: Intrinsic parameter redundancies for capture-recapture-recovery $x^1/y/(z_1;z_2)$ models with first-year survival probabilities (Table 3)

Model	Rank	Deficiency	Method of Proof
T^1/T/(C;C)	n_1+n_2+1	0	Full Rank Theorem
T^1/T/(C;T)	n_1+2n_2	0	Full Rank Theorem
T^1/T/(C;A)	n_1+2n_2	0	Full Rank Theorem
T^1/T/(C;A,T)	$E+n_1+n_2$	0	Full Rank Theorem
T^1/T/(T;C)	n_1+2n_2	0	Extension Theorem
T^1/T/(T;T)†	n_1+3n_2-1	0	Extension Theorem ($d=0$†)
T^1/T/(T;A)	n_1+3n_2-1	0	Extension Theorem
T^1/T/(T;A,T)	$E+n_1+2n_2-2$	1	Reparameterisation
T^1/T/(A;C)	n_1+2n_2	0	Full Rank Theorem
T^1/T/(A;T)	n_1+3n_2-1	0	Full Rank Theorem
T^1/T/(A;A)	n_1+3n_2-1	0	Full Rank Theorem
T^1/T/(A;A,T)	$E+n_1+2n_2-1$	0	Full Rank Theorem
T^1/T/(A,T;C)	$E+n_1+n_2$	0	Extension Theorem
T^1/T/(A,T;T)	$E+n_1+2n_2-2$	1	Reparameterisation
T^1/T/(A,T;A)	$E+n_1+2n_2-1$	0	Extension Theorem
T^1/T/(A,T;A,T)‡	$2E+n_1+n_2-3$	2	Reparameterisation
T^1/A/(C;C)	n_1+n_2+1	0	Full Rank Theorem
T^1/A/(C;T)	n_1+2n_2	0	Full Rank Theorem
T^1/A/(C;A)	n_1+2n_2	0	Full Rank Theorem
T^1/A/(C;A,T)	$E+n_1+n_2$	0	Full Rank Theorem
T^1/A/(T;C)	n_1+2n_2	0	Full Rank Theorem
T^1/A/(T;T)	n_1+3n_2-1	0	Full Rank Theorem
T^1/A/(T;A)	n_1+3n_2-1	0	Full Rank Theorem
T^1/A/(T;A,T)	$E+n_1+2n_2-1$	0	Full Rank Theorem
T^1/A/(A;C)	n_1+2n_2	0	Extension Theorem
T^1/A/(A;T)	n_1+3n_2-1	0	Extension Theorem
T^1/A/(A;A)	n_1+3n_2-2	1	Reparameterisation
T^1/A/(A;A,T)	$E+n_1+2n_2-2$	1	Reparameterisation
T^1/A/(A,T;C)	$E+n_1+n_2$	0	Extension Theorem
T^1/A/(A,T;T)	$E+n_1+2n_2-1$	0	Extension Theorem
T^1/A/(A,T;A)	$E+n_1+2n_2-2$	1	Reparameterisation
T^1/A/(A,T;A,T)‡	$2E+n_1+n_2-3$	2	Reparameterisation
T^1/A,T/$(z_1;z_2)$	Identical to A,T/$(z_1;z_2)$ models: See Table 4.3b		

†: when $n_2>n_1$, rank decreases by 1 and parameter deficiency increases by 1;
‡: when $n_2>n_1$, rank increases by 1 and parameter deficiency decreases by 1;
$E=n_1n_2-\frac{1}{2}n_1^2+\frac{1}{2}n_1$.

is equal to $0.2^8 = 0.00000256$ and the probability of the life-history 11112 is equal to $0.8 \times 0.2^7 = 0.00001024$, both very small probabilities. This section is similar to our analyses performed in Section 3.7 for capture-recapture models where we viewed imperfect data sets and show how this affects parameter estimation for a variety of different parameter dependent models. In Example 4.6 below we show how we can use `Maple` code to obtain extrinsic parameter redundancy results.

Example 4.6 - Extrinsic parameter redundancy in the capture-recapture-recovery model: [See electronic appendix `example4.6.mw`] Consider the A,T/(C;T) capture-recapture-recovery model where the model has age- and time-dependent survival probabilities, constant recapture probabilities, and time-dependent recovery probabilities for a study where $n_1 = 2$ and $n_2 = 3$. Suppose we observed only the distinct life-histories 1200, 1002, 1111, 1100, 1110, 1001, 0102, 0111 and 0100. These life-histories are given in the matrix **Data** with each row representing a different life-histories and each column corresponding to each capture/recovery occasion during the study:

$$\mathbf{Data} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The `Maple` procedure `caprecaprecovhistories` converts the life-histories above into the exhaustive terms given in $\boldsymbol{\kappa}$ below

```
kappa := caprecaprecovhistories(Data,4,1,2);
# Inputs: (Data,y,z1,z2); Data = Data of all life-histories;
# y = survival probability; z1 = recapture probability;
# z2 = recovery probability;
# for y, z1 and z2: '1'=C, '2'=T, '3'=A, '4'=A,T.
```

This generates the model's exhaustive summary as

$$\boldsymbol{\kappa} = \begin{bmatrix} (1-\phi_{1,1})\lambda_1 \\ \phi_{1,1}(1-p)\phi_{2,2}(1-p)(1-\phi_{3,3})\lambda_3 \\ \phi_{1,1}p\phi_{2,2}p\phi_{3,3}p \\ \phi_{1,1}p\left[(1-\phi_{2,2})(1-\lambda_2)+\phi_{2,2}(1-p)\left\{(1-\phi_{3,3})(1-\lambda_3)+\phi_{3,3}(1-p)\right\}\right] \\ \phi_{1,1}p\phi_{2,2}p\left\{(1-\phi_{3,3})(1-\lambda_3)+\phi_{3,3}(1-p)\right\} \\ \phi_{1,1}(1-p)\phi_{2,2}(1-p)\phi_{3,3}p \\ \phi_{1,2}(1-p)(1-\phi_{2,3})\lambda_3 \\ \phi_{1,2}p\phi_{2,3}p \\ (1-\phi_{1,2})(1-\lambda_2)+\phi_{1,2}(1-p)\left\{(1-\phi_{2,3})(1-\lambda_3)+\phi_{2,3}(1-p)\right\} \end{bmatrix}.$$

The rank and parameter deficiency of the model can then be found as described in Section 2.2:

```
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),theta):
r := Rank(DD); d := Dimension(theta)-r;
```

The parameters in this model are $\boldsymbol{\theta} = [\phi_{1,1},\ \phi_{1,2},\ \phi_{2,2},\ \phi_{2,3},\ \phi_{3,3},\ p,\ \lambda_1,\ \lambda_2,\ \lambda_3]^T$. The derivative matrix is given in the `Maple` file `example4.6.mw` only to save space. This derivative matrix has the model rank and parameter deficiency as

$$\mathbf{r} := \mathbf{9}, \mathbf{d} := \mathbf{0}.$$

This shows that this small data set has no parameter redundancy, even if the survival parameters are age- and time-dependent in the A,T/(C;T) capture-recapture-recovery model. □

Example 4.7 - The cormorants data set for the capture-recapture-recovery model: This data set examines Great Cormorants (*Phalacrocorax carbo sinensis*) between 1981 and 1993 from Hènaux et al. (2007). Six different colonies were observed in the study and recapture/recovery data were observed for each animal of each colony. The data was formed through the observation of cormorants which were marked by the use of coloured rings, and these cormorants were then observed using a strong telescope. The data consists of 14,018 individual cormorant life-histories of which 1,939 cormorants were recovered dead. The number of first-capture occasions and recapture/recovery occasions is equal to twelve in this study.

We consider three cases of colonies for our analysis: One case is where only the life-histories of the cormorants from colony 3 are considered, another case is where only the life-histories of the cormorants from colony 1 are considered, and a final case is where all of the life-histories in every colony are considered. The colony 3 only data set is very sparse and only has 121 distinct life-histories were observed out of a possible 12,282 distinct life-histories that a perfect data set would have. The colony 1 only data has 465 distinct life-histories and the data for all of the colonies has 580 distinct life-histories. More detail on the study design can be found in Hènaux et al. (2007). The cormorant data are typically analysed using multi-site models, see for example Borysiewicz (2008) and McCrea et al. (2010, 2012). We use our capture-recapture-recovery models to analyse this data set for illustrative purposes only.

We examine the three different selections of colonies by analysing all of the different parameter dependency possibilities and show how sparse data can affect the ability to estimate parameters. This is a similar analysis to that of the Dippers data set of Marzolin (1988) in Section 3.7. We use the `Maple` procedure `caprecaprecovhistories` from Example 4.6 to generate these extrinsic parameter redundancy results shown in Tables 4.6a and 4.6b. The second column in the following tables displays the number of parameters in model. If some of the model parameters do not appear at all, they are excluded from the count of parameters so that these do not contribute to the parameter deficiency. As there is no first-capture in the 11th capture occasion of this study the parameters $\phi_{1,12}$, $p_{2,13}$ and $\lambda_{1,12}$, if the parameter dependencies are age- and time-dependent, will not appear in any exhaustive summary terms. For colony 3, this problem is extended where there is no first-capturing between the 8th and 11th capture occasions of the study. This is taken into account as it changes the number of parameters in some of the models and we use m to denote this in Tables 4.6a and 4.6b. The third column shows the parameter deficiency for the colony 3 only data set, the fourth column for the colony 1 only data set, and the fifth column for the data set that contains all colonies, as well as its intrinsic parameter redundancy shown in the final column.

The colony 3 data set is extremely sparse and this results in most models being parameter redundant. The other data sets though still have most models remaining full rank, even with the relative sparseness of the data, e.g. for the colony 1 data set, only 3.8% of the possible life-histories are observed in the data (465 out of the 12,282). The exceptions are models where at least one parameter is age- and time-dependent as it can be quite hard to estimate every parameter in these cases. □

Table 4.6a: Extrinsic parameter redundancy results for capture-recapture-recovery $y/(z_1; z_2)$ models for the cormorants data set of Hènaux et al. (2007) (Table 1)

Model	Number of parameters	Parameter deficiency of			
		Colony 3	Colony 1	All colonies	Intrinsic model
C/(C;C)	3	0	0	0	0
C/(C;T)	14	0	0	0	0
C/(C;A)	14	1	0	0	0
C/(C;A,T)	$80 - m$	4	0	0	0
C/(T;C)	14	0	0	0	0
C/(T;T)	25	1	0	0	0
C/(T;A)	25	1	0	0	0
C/(T;A,T)	$91 - m$	5	0	0	0
C/(A;C)	14	1	0	0	0
C/(A;T)	25	1	0	0	0
C/(A;A)	25	2	0	0	0
C/(A;A,T)	$91 - m$	7	0	0	0
C/(A,T;C)	80	8	0	0	0
C/(A,T;T)	91	9	0	0	0
C/(A,T;A)	91	11	0	0	0
C/(A,T;A,T)	$157 - m$	36	1	1	0
T/(C;C)	14	0	0	0	0
T/(C;T)	25	1	0	0	0
T/(C;A)	25	1	0	0	0
T/(C;A,T)	$91 - m$	5	0	0	0
T/(T;C)	25	1	0	0	0
T/(T;T)	36	5	1	1	1
T/(T;A)	36	1	0	0	0
T/(T;A,T)	$102 - m$	8	1	1	1
T/(A;C)	25	1	0	0	0
T/(A;T)	36	1	0	0	0
T/(A;A)	36	2	0	0	0
T/(A;A,T)	$102 - m$	9	0	0	0
T/(A,T;C)	$91 - m$	9	0	0	0
T/(A,T;T)	$102 - m$	12	1	1	1
T/(A,T;A)	$102 - m$	12	0	0	0
T/(A,T;A,T)	$168 - 2m$	41	2	2	1
$m = 10$ for colony 3, $m = 1$ for colony 1 and all colonies, and $m = 0$ for intrinsic.					

Table 4.6b: Extrinsic parameter redundancy results for capture-recapture-recovery $y/(z_1; z_2)$ models for the cormorants data set of Hènaux et al. (2007) (Table 2)

Model	Number of parameters	Parameter deficiency of			
		Colony 3	Colony 1	all colonies	Intrinsic model
A/(C;C)	14	1	0	0	0
A/(C;T)	25	1	0	0	0
A/(C;A)	25	2	0	0	0
A/(C;A,T)	$91 - m$	8	0	0	0
A/(T;C)	25	1	0	0	0
A/(T;T)	36	1	0	0	0
A/(T;A)	36	2	0	0	0
A/(T;A,T)	$102 - m$	9	0	0	0
A/(A;C)	25	2	0	0	0
A/(A;T)	36	2	0	0	0
A/(A;A)	36	6	2	2	1
A/(A;A,T)	$102 - m$	13	2	2	1
A/(A,T;C)	$91 - m$	12	0	0	0
A/(A,T;T)	$102 - m$	13	0	0	0
A/(A,T;A)	$102 - m$	17	1	1	1
A/(A,T;A,T)	$168 - 2m$	43	2	2	1
A,T/(C;C)	$80 - m$	3	0	0	0
A,T/(C;T)	$91 - m$	5	0	0	0
A,T/(C;A)	$91 - m$	7	0	0	0
A,T/(C;A,T)	$157 - 2m$	42	10	6	0
A,T/(T;C)	$91 - m$	6	0	0	0
A,T/(T;T)	$102 - m$	9	0	0	0
A,T/(T;A)	$102 - m$	10	1	1	0
A,T/(T;A,T)	$168 - 2m$	48	12	8	1
A,T/(A;C)	$91 - m$	8	0	0	0
A,T/(A;T)	$102 - m$	10	0	0	0
A,T/(A;A)	$102 - m$	12	1	1	1
A,T/(A;A,T)	$168 - 2m$	49	12	8	1
A,T/(A,T;C)	$157 - 2m$	41	7	6	0
A,T/(A,T;T)	$168 - 2m$	46	8	7	1
A,T/(A,T;A)	$168 - 2m$	46	8	7	1
A,T/(A,T;A,T)	$234 - 3m$	96	32	28	12
$m = 10$ for colony 3, $m = 1$ for colony 1 and all colonies, and $m = 0$ for intrinsic.					

As is the case for the capture-recapture model, we have created the `Maple` procedure `fullhistcaprecaprecov` to list all the life-histories for a complete data set to check the intrinsic results from Tables 4.3a and 4.3b. This code is more complex than the code for the capture-recapture model as while there is the extra possibility of being recovered dead, there can only be a single '2' observation in each life-history at most. This `Maple` code allows the checking of intrinsic parameter redundancy results for the capture-recapture-recovery model for certain n_1 and n_2 values. This is demonstrated in Example 4.6 below.

Example 4.6 revisited: [See electronic appendix `example4.6.mw`] Consider a study where there were five first-capture occasions and six recapture/recovery occasions when we are viewing the A,T/(A;A) model for this example. All possible distinct life-histories are generated using the `Maple` code

```
Data := fullhistcaprecaprecov(5,6);
# Inputs: (n1,n2);
# n1 = number of first-capture occasions;
# n2 = number of recapture/recovery occasions.
```

This code gives the matrix

$$\mathbf{Data} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 & 2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

which contains all the possible distinct life-histories. The output from this code does contain a number of rows which have zeros for every entry: This is due to using the `combinat` package in `Maple` which means that not all rows are actually valid life-histories, and we replace these with rows containing zeros for all its entries. This allows us to easily exclude these invalid life-histories as these rows of zeros display the result '0' when the `Maple` procedure `caprecaprecovhistories` is performed on these rows and the rows do not contribute any exhaustive summary terms in $\boldsymbol{\kappa}$. To obtain the

final rank and parameter redundancy, we use the usual procedures of Section 2.2 as given below:

```
kappa := caprecaprecovhistories(Data,4,3,3):
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

The exhaustive summary, parameter set and derivative matrix are not given here to save space but can be found in the `Maple` file `example4.6.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. These life-histories have the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{31}, \mathbf{d} := \mathbf{1}$.

This agrees with the result from Table 4.3b where the rank is $r = E + 2n_2 - 1 = 20 + (2 \times 6) - 1 = 31$ and the parameter deficiency is $d = 1$. High values of n_1 and n_2 can produce a computationally slow procedure though we have used the hybrid symbolic-numerical approach of Choquet and Cole (2012) and the logvector form of the exhaustive summary to speed up computation. Note that this code provides a verification of the general results from Tables 4.3a and 4.3b and is not a proof of the results. □

4.7 An extrinsic sparseness statistic for capture-recapture-recovery models

We extend our analysis of extrinsic parameter redundancy in capture-recapture-recovery models by generalising which possible life-histories a practical data set may have. This is similar to the analysis performed in Section 3.8 where we generalised capture-recapture data sets. We define a statistic, c, to indicate the maximum difference between the first capture occasion and the last recapture/recovery occasion, where $1 \leq c < n_2$. If we suppose that we have all life-histories where the difference between first capture and last recapture or recovery is equal to c, then we can calculate the parameter deficiency only using the probabilities of those life-histories. This is similar to the methodology in Section 3.8 apart from the fact we are also considering the life-histories which are recovered dead up to c recovery occasions later as well as live recaptures. Real data will rarely have this exact pattern of life-histories, but we

would expect a data set which is sparse with a lower number of distinct life-histories to behave like a model with a low value of c.

4.7a and 4.7b show the parameter deficiency of capture-recapture-recovery models, given a certain number of recapture/recovery occasions n_2, with a maximum difference c between first capture occasion and the last recapture/recovery occasion. We consider $n_1 = n_2$ only as consistent patterns for parameter deficiencies do not happen if this is not the case. These results were generated by the hybrid symbolic-numerical approach from Choquet and Cole (2012) to obtain a model's parameter deficiency instead of the pure symbolic method due to computational complexity. The parameter deficiency of the model when $c > 1$ is shown in the second column, the parameter deficiency of the model when $c = 1$ is shown in the third column, with the model's intrinsic parameter deficiency in the final column. We also demonstrate how the parameter redundancy can be obtained in Example 4.8 below.

Example 4.8: Consider the case where the difference between first capture occasion and last recapture/recovery occasion is a maximum of $c = 2$. A model with four first-capture and recapture/recovery occasions with $c = 2$ has the life-histories {11100, 10100, 11000, 10000, 01110, 01010, 01100, 01000, 00111, 00101, 00110, 00100, 00011, 00010, 12000, 11200, 10200, 01200, 01120, 01020, 00120, 00112, 00102, 00012}, and not the rest of the possible life-histories as they all have either three or four occasions between first capture and last recapture/recovery. We can then check if a certain capture-recapture-recovery model has parameter redundancy from Tables 4.7a and 4.7b. In this case where $c = 2$ and $n_1 = n_2 = 4$, a total of 19 different $y/(z_1; z_2)$ models are not parameter redundant. If we observe more life-histories so that $c = 3$, then a further 11 $y/(z_1; z_2)$ models also become full rank. □

4.8 Fidelity in the capture-recapture-recovery model

The final sections in this chapter considers a capture-recapture-recovery model involving additional parameters regarding an animal's **fidelity**. Fidelity in this context means the probability that an animal remains at the sample location during the study. This in turn means the probability of animal fidelity is the complement of the probability that the animal emigrates. We only consider permanent emigration in this thesis where the animal emigrates from the sample location with no probability of return to the site, but an alternative model could be considered where there is some probability of immigration back into the location. In our model, while it is possible to recover an

Table 4.7a: Parameter redundancies for capture-recapture-recovery $y/(z_1;z_2)$ models where there is a maximum difference c between the number of occasions between first capture and last recapture/recovery (Table 1)

Model	Deficiency when $c>1$	Deficiency when $c=1$	Intrinsic
C/(C;C)	0	0	0
C/(C;T)	0	0	0
C/(C;A)	0	0	0
C/(C;A,T)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)$	0
C/(T;C)	0	0	0
C/(T;T)	0	0	0
C/(T;A)	0	1	0
C/(T;A,T)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)$	0
C/(A;C)	0	0	0
C/(A;T)	0	0	0
C/(A;A)	n_2-c	n_2-1	0
C/(A;A,T)	$\frac{1}{2}(n_2-c+1)(n_2-c)$	$\frac{1}{2}n_2(n_2-1)+1$	0
C/(A,T;C)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)$	0
C/(A,T;T)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)$	0
C/(A,T;A)	$\frac{1}{2}(n_2-c+1)(n_2-c)$	$\frac{1}{2}n_2(n_2-1)+1$	0
C/(A,T;A,T)	$(n_2-c)^2$	$(n_2-1)^2+1$	0
T/(C;C)	0	0	0
T/(C;T)	0	0	0
T/(C;A)	0	0	0
T/(C;A,T)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)+1$	0
T/(T;C)	0	0	0
T/(T;T)	1	1	1
T/(T;A)	0	1	0
T/(T;A,T)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)$	1
T/(A;C)	0	0	0
T/(A;T)	0	1	0
T/(A;A)	n_2-c	n_2-1	0
T/(A;A,T)	$\frac{1}{2}(n_2-c+1)(n_2-c)$	$\frac{1}{2}n_2(n_2-1)+1$	0
T/(A,T;C)	$\frac{1}{2}(n_2-c)(n_2-c-1)$	$\frac{1}{2}(n_2-1)(n_2-2)+1$	0
T/(A,T;T)	$\frac{1}{2}(n_2-c)(n_2-c-1)+1$	$\frac{1}{2}(n_2-1)(n_2-2)$	1
T/(A,T;A)	$\frac{1}{2}(n_2-c+1)(n_2-c)$	$\frac{1}{2}n_2(n_2-1)+1$	0
T/(A,T;A,T)	$n_2^2-2n_2c+c^2+1$	$n_2^2-n_2+1$	1

Table 4.7b: Parameter redundancies for capture-recapture-recovery $y/(z_1; z_2)$ models where there is a maximum difference c between the number of occasions between first capture and last recapture/recovery (Table 2)

Model	Deficiency when $c > 1$	Deficiency when $c = 1$	Intrinsic
A/(C;C)	0	0	0
A/(C;T)	0	0	0
A/(C;A)	$n_2 - c$	$n_2 - 1$	0
A/(C;A,T)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1)$	0
A/(T;C)	0	1	0
A/(T;T)	0	1	0
A/(T;A)	$n_2 - c$	n_2	0
A/(T;A,T)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1) + 1$	0
A/(A;C)	$n_2 - c$	$n_2 - 1$	0
A/(A;T)	$n_2 - c$	$n_2 - 1$	0
A/(A;A)	$2(n_2 - c)$	$2n_2 - 2$	1
A/(A;A,T)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + n_2 - c$	$\frac{1}{2}n_2(n_2 - 1)$	1
A/(A,T;C)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1) + 1$	0
A/(A,T;C)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1) + 1$	0
A/(A,T;A)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + n_2 - c + 1$	$\frac{1}{2}n_2(n_2 - 1) + 1$	1
A/(A,T;A,T)	$3(n_2 - c) + (n_2 - c - 1)^2$	$n_2^2 - n_2 + 1$	1
A,T/(C;C)	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - 1)(n_2 - 2)$	0
A,T/(C;C)	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - 1)(n_2 - 2) + 1$	0
A,T/(C;A)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1)$	0
A,T/(C;A,T)	$(n_2 - c)^2$	$(n_2 - 1)^2 + 1$	0
A,T/(T;C)	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - 1)(n_2 - 2) + 1$	0
A,T/(T;T)	$\frac{1}{2}(n_2 - c)(n_2 - c - 1)$	$\frac{1}{2}(n_2 - 1)(n_2 - 2) + 1$	0
A,T/(T;A)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1) + 1$	0
A,T/(T;A,T)	$n_2^2 - 2n_2c + c^2 + 1$	$n_2^2 - n_2 + 1$	1
A,T/(A;C)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1)$	0
A,T/(A;T)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c)$	$\frac{1}{2}n_2(n_2 - 1) + 1$	0
A,T/(A;A)	$\frac{1}{2}(n_2 - c + 1)(n_2 - c) + n_2 - c$	$\frac{1}{2}n_2(n_2 - 1)$	1
A,T/(A;A,T)	$3(n_2 - c) + (n_2 - c - 1)^2$	$n_2^2 - n_2 + 1$	1
A,T/(A,T;C)	$(n_2 - c)^2$	$(n_2 - 1)^2 + 1$	0
A,T/(A,T;T)	$n_2^2 - 2n_2c + c^2 + 1$	$n_2^2 - n_2 + 1$	1
A,T/(A,T;A)	$3(n_2 - c) + (n_2 - c - 1)^2$	$n_2^2 - n_2 + 1$	1
A,T/(A,T;A,T)	$\frac{3}{2}(n_2 - c + 1)(n_2 - c) + c$	$\frac{1}{2}3n_2(n_2 - 1) + 1$	n_1

animal dead at any place, the animal must not have permanently emigrated to be able to be recaptured alive. This adds complication to the model but allows us to estimate animal fidelity probabilities, as well as animal survival probabilities. This **fidelity model** is given in Burnham (1993, Section 3.2) and can be fit using the software package `MARK` can be seen in Cooch and White (2014, Chapter 9). Examples of the use of fidelity models in statistical ecology include Webb et al. (2010) for white-tailed deer and Collier et al. (2012) for white-winged doves in Texas.

We denote fidelity, F_j, as the probability that an animal remains in the sample location (study area) during year j, for $j = i, \ldots, n_2$. It is important to distinguish between *apparent* and *real* survival probabilities, as mentioned before in Section 3.1. Previously in this thesis, we do not change notation between apparent and real survival probabilities and use the notation of ϕ for all survival parameters. However, this distinction is made more clearly in this fidelity model due to an animal possibly emigrating from the study location. We if let ϕ_j denote the time-dependent probability of apparent animal survival, conditional on the animal remaining in the study location, and S_j denote the time-dependent real animal survival probability, then

$$S_j = \frac{\phi_j}{F_j}, \tag{4.7}$$

holds. If there is no permanent emigration then $F_j = 1$, so that $S_j = \phi_j$, and the parameters of apparent and real survival are the same as every animal remains in the study location. If some of the animals emigrate during the study so that $F_j < 1$, while it is still possible to recover these animals dead, it is *not* possible to recapture those animals alive, so apparent and real survival parameters are then different. For the fidelity model in this thesis, we use the real animal survival parameters S_j instead of the apparent survival parameters as this is consistent with the notation from Burnham (1993) and Cooch and White (2014).

We let n_1 denote the number of first-capture occasions and n_2 denote the number of recapture occasions in the study where $n_1 \leq n_2$. If we consider a study with yearly intervals say, then there is one capture occasion and one recovery occasion per year. Consequently, while there are still n_2 recapture occasions during the study, there are now $n_2 + 1$ recovery occasions, as an animal can be recovered dead if the animal was first-captured at the beginning of the study and recovered dead before the first recapture occasion, along with there being a recovery occasion after the final recapture occasion in the study. While the previous model from Section 4.2 does not consider the

possibility that we recapture the animal alive and recover the animal in consecutive recapture and recovery occasions during a single year, we now include this possibility in this model. The reason why this was not considered in the model of Section 4.2 is because the data for the previous capture-recapture-recovery model generally does not consider this as a possibility of being recorded, while data coming from the analyses using fidelity models does include this possibility. We also assume that generally we have no information from where the dead recovery was made, however this could be factored into a different fidelity model where this information can give some indication on whether those dead animals migrated or not.

The notation for the fidelity model used in Burnham (1993) and Cooch and White (2014) uses a code of '1' and '0' values in pairs of observations. The pairs of numbers represent if they recaptured the animal alive and if they recovered the animal dead during a single recapture/recovery occasion. We however use a different notation here to keep all of our notation consistent with the rest of this chapter. We still use '1' to denotes a live recapture and '2' to represent a dead recovery, with '0' to denote that an animal is not recaptured or recovered at that occasion. We then introduce a '**3**' to denote an animal which was recaptured alive at recapture occasion j *and* recovered dead between j and $j+1$. This possibility of a '3' entry is a special case in this model which we did not consider in the previous model of Section 4.2, where both a live recapture and dead recovery are possible simultaneously in a single time frame. Note that our life-history notation is also the same as the binary representation of each pair of numbers from the notation given in Burnham (1993) and Cooch and White (2014) if the digits in each pair are switched, e.g. the probability of the life-history 10 10 10 11 has the same expression as our probability of the life-history 1113. In this fidelity model, as in the previous models in this thesis, there are n_2+1 digits in the life-history code as we do not consider the first capture occasion as a recapture occasion. We illustrate some possible life-histories in Examples 4.9 to 4.13 below.

Example 4.9: Consider the life-history for a study of three recapture and four recovery occasions where an animal was first-captured at the first capture occasion. The animal was then not recaptured during the first recapture occasion, recaptured alive during the second recapture occasion, and was finally being recaptured alive and recovered dead in the last capture and recovery occasions of the study. This would have the life-history code 1013, where the first 1 indicates the initial capture, the 0 indicates no recapture, the next 1 indicates an alive recapture, and the 3 indicates the animal was recaptured alive before it was also recovered dead in the same year. In the notation

from Burnham (1993) and Cooch and White (2014) the life-history would alternatively be expressed as 10 00 10 11. □

Example 4.10: Consider a very similar animal life-history, where instead of being both recaptured alive and recovered dead in the last capture and recovery occasions of the study, the animal is only recovered dead and *not* recaptured alive. This would have the life-history code 1012, where the first 1 indicates the initial capture, the 0 indicates no recapture, the next 1 indicates an alive recapture, and the 2 indicates the animal was only recovered dead. In the notation from Burnham (1993) and Cooch and White (2014) the life-history would alternatively be expressed as 10 00 10 01. □

Example 4.11: It is possible to have no live recaptures observed for an animal, while it is still recovered dead at the final recovery occasion. For a study of four capture and three recapture occasions, this would have the life-history code 1002, where Burnham (1993) and Cooch and White (2014) would alternatively express this life-history as 10 00 00 01. As we will demonstrate later on in this section, the probability of this life-history is complex as while the animal is recovered dead at the end of the study, there is no information on whether the animal emigrated from the location study or not before its recovery dead. □

Example 4.12: There can be life-histories where the animal is not recovered dead at any point. For example, take the life-history where the animal is first-captured at the first capture occasion, recaptured alive at the second and third capture occasions, but not recapture alive at the fourth capture occasion. It is also not recovered dead at any point in the study. This would have the life-history code 1110, where Burnham (1993) and Cooch and White (2014) would alternatively express this life-history as 10 10 10 00. A χ term will be used to express the uncertainty of whether the animal is alive or dead at the final recapture and recovery occasions of the study, as we shall demonstrate later on in this section. □

Example 4.13: Our final example would be one possibly that an animal is recovered dead straight after the animal's first-capture occasion in the study. This would have the life-history code as 3000, where Burnham (1993) and Cooch and White (2014) would alternatively express this life-history as 11 00 00 00. □

Let p_j denote the probability that an animal is recaptured at capture occasion j (i.e. recapture occasion $j-1$), for $j = 2, \ldots, n_2+1$ and λ_j denote the probability that

an animal is recovered dead in the jth year of the study, for $j = 1, \ldots, n_2 + 1$. There are a total of $n_2 + 1$ λ_j parameters as it is possible to be recovered dead before the first recapture occasion as well as after the final recapture occasion in the study. We let S_j denote the probability that an animal at time j survives until time $j + 1$ given that the animal has survived up to that point, for $j = 1, \ldots, n_2 + 1$, and F_j denote the probability of animal fidelity (where $1 - F_j$ is the probability of permanent animal emigration), that an animal at time j remains in the study location until capture occasion j, for $j = 1, \ldots, n_2$. We do not consider any age-dependent parameters as age is rarely explored in the data sets which are fitted using capture-recapture-recovery fidelity models. We now revisit Examples 4.9 to 4.13 below and demonstrate how the probabilities of those life-histories can be obtained.

Example 4.9 revisited: In the life-history 1013, the animal was recaptured alive at each recapture occasion apart from the first recapture occasion, and was also recovered dead at the final recovery occasion. If we assume that all parameters are time-dependent, the probability of this life-history is

$$Pr(1013) = S_1F_1(1 - p_2)S_2F_2p_3S_3F_3p_4(1 - S_4)\lambda_4.$$

This shows an example where an animal was recaptured alive and recovered dead in the final occasions of the study. It is also assumed to have not permanently emigrated during the study as the animal was recaptured alive in the final recapture occasion of its life. □

Example 4.10 revisited: A similar life-history is 1012 where the animal was *not* recaptured alive during the last recapture occasion when the animal was then recovered dead. If we assume that all parameters are time-dependent, the probability of this life-history is

$$Pr(1012) = S_1F_1(1 - p_2)S_2F_2p_3S_3\left\{(1 - F_3) + F_3(1 - p_4)\right\}(1 - S_4)\lambda_4.$$

The added complication here stems for the fact that while we know the animal survived until the final recovery occasion, we are unclear whether the animal emigrated during this time or the animal remained at the location but was not recaptured. We deal with this complication when we are unsure that the animal has emigrated or not after the third capture occasions by using a γ term to represent $\gamma_{3,4} = \{(1 - F_3) + F_3(1 - p_4)\}$. We give the general form for γ in Equation (4.10) later in this section. □

Example 4.11 revisited: This γ term seen in Example 4.10, which represents our uncertainty over whether or not the animal has emigrated or not before a recovery dead, is more complex when there is a greater number of years between last capture and recovery dead. If we observe the life-history 1002 where the animal was never recaptured alive but was recovered dead at the final recovery occasion of the study, then if we assume that all parameters are time-dependent, the probability of this life-history is

$$\begin{aligned} Pr(1002) &= S_1S_2S_3\left\{(1-F_1)+F_1(1-p_2)\left[(1-F_2)\right.\right. \\ &\quad \left.\left.+\ F_2(1-p_3)\left\{(1-F_3)+F_3(1-p_4)\right\}\right]\right\}(1-S_4)\lambda_4 \\ &= S_1S_2S_3\gamma_{1,4}(1-S_4)\lambda_4. \end{aligned}$$

The complication here is the fact that while we know the animal survived until the final recovery occasion as we recovered dead at that time, we do not know if the animal has permanently emigrated or not. Furthermore, this emigration could have happened at any time before its recovery, thus the long expression for $\gamma_{1,4}$ in this case as we are unsure about when the animal emigrated if at all. □

Example 4.12 revisited: We use χ terms in our capture-recapture-recovery models of Section 4.2 to represent the fact we are uncertain whether the animal has died and not been recovered, or is still alive but has not been recaptured alive. The χ terms are also used in the fidelity model, such as in the animal life-history 1110 where the animal was recaptured alive during the first and second recapture occasions but we are uncertain of its survival after the second recapture occasion. However, this χ term is more complex in the fidelity model as we also have to consider animal emigration as well. If we assume that all parameters are time-dependent, the probability of this life-history is

$$\begin{aligned} Pr(1110) &= S_1F_1p_2S_2F_2p_3\left[(1-S_3)(1-\lambda_3)+S_3(1-F_3)\left\{(1-S_4)(1-\lambda_4)\right.\right. \\ &\quad \left.\left.+\ S_4\right\}+S_3F_3(1-p_4)\left\{(1-S_4)(1-\lambda_4)+S_4\right\}\right] \\ &= S_1F_1p_2S_2F_2p_3\chi_3. \end{aligned}$$

This χ_3 term represents that we are uncertain about the animal after the second recapture occasion of the study: The animal could either have died just after the second recapture occasion and not been recovered with probability $(1-S_3)(1-\lambda_3)$, the animal could have emigrated and then either died with no recovery or survived with

probability $S_3(1-F_3)\{(1-S_4)(1-\lambda_4)+S_4\}$, or the animal could have remained in the study location and was not captured ar capture occasion four (recapture occasion three) and then either died with no recovery or survived with probability $S_3F_3(1-p_4)\{(1-S_4)(1-\lambda_4)+S_4\}$. This χ term can then have a very wide range of possibilities if more '0' observations are in the life-history, say the life-history 1000 for example. We give the general form for χ in Equation (4.8) later in this section. □

Example 4.13 revisited: Finally, we can also observe a dead recovery straight after first-capture in this model, denoted by the life-history 3000. If we assume that all parameters are time-dependent, the probability of this life-history is simply

$$Pr(3000) = (1-S_1)\lambda_1.$$

□

A model for these life-histories can be seen in the likelihood given in Burnham (1993, pp.205). We however would like to examine individual life-histories as a base for generating an exhaustive summary. We can write down the probability of a particular life-history occurring as we did in for the previous capture-recapture-recovery model in Section 4.2, but it is more complex here as permanent emigration has to also be considered. It is also true when an animal permanently emigrates that the animal cannot be recaptured alive it could possibly still be recovered dead. Let

$$\chi_j = (1-S_j)(1-\lambda_j) + S_j(1-F_j)\Psi_{j+1} + S_jF_j(1-p_{j+1})\chi_{j+1}, \tag{4.8}$$

where $\chi_{n_2+1} = (1-S_{n_2+1})(1-\lambda_{n_2+1}) + S_{n_2+1}$, and

$$\Psi_j = (1-S_j)(1-\lambda_j) + S_j\Psi_{j+1}, \tag{4.9}$$

with $\Psi_{n_2+1} = (1-S_{n_2+1})(1-\lambda_{n_2+1}) + S_{n_2+1}$. The term Ψ_j denotes the probability of either being not recovered dead or surviving, given that the animal has permanently emigrated after the last time the animal was recaptured. This complexity in the χ term is encompassed in Burnham (1993) using his notations of λ_d and λ_l.

Suppose an animal was last recaptured alive at capture occasion w and was also recovered dead at time b, then let the probability of not capturing the animal alive between time w and the time the animal was recovered dead at time b be

$$\gamma_{w,b} = (1 - F_w) + F_w(1 - p_{w+1})\gamma_{w+1,b}, \tag{4.10}$$

where $\gamma_{b,b} = 1$. This is a necessary term for the case where the animal is recovered dead and not recaptured alive at the same time. The complication in the γ term is because an animal could emigrate or not but still be recovered dead at a later point in the study.

Let δ_k denote the life-history entry at time k. The possible individual life-histories at a specific time point are $\delta_k = 3$ if the animal is both recaptured alive and recovered dead at occasion k, $\delta_k = 2$ if the animal is only recovered dead at occasion k, $\delta_k = 1$ if the animal is only recaptured alive in year k, and $\delta_k = 0$ if the animal is neither recaptured or recovered at occasion k. The life-history entry δ_1 can only be '1' or '3' as the animal cannot be recaptured alive until the next capture occasion at entry δ_2. Let an animal be first-captured at time a and last seen either by recapturing or recovery at time b. If we further let w denote the last capture occasion the animal was recaptured alive, the probability associated with a particular life-history, h_X, is equal to

$$Pr(h_X) = \begin{cases} \displaystyle\prod_{k=a+1}^{b} S_{k-1}F_{k-1}\left\{\delta_k p_k + (1-\delta_k)(1-p_k)\right\}\chi_b & \text{if } \delta_b = 1, \\ \displaystyle\prod_{k=a+1}^{w} S_{k-1}F_{k-1}\left\{\delta_k p_k + (1-\delta_k)(1-p_k)\right\} & \text{if } \delta_b = 2, \\ \qquad \times \displaystyle\prod_{m=w}^{b-1} S_m\gamma_{w,b}(1-S_b)\lambda_b & \\ \displaystyle\prod_{k=a+1}^{b-1} S_{k-1}F_{k-1}\left\{\delta_k p_k + (1-\delta_k)(1-p_k)\right\} & \text{if } \delta_b = 3. \\ \qquad \times S_{b-1}F_{b-1}p_b(1-S_b)\lambda_b & \end{cases} \tag{4.11}$$

An overall likelihood can be obtained as the product of the N different life-histories observed as

$$L = \prod_{X=1}^{N} Pr(h_X).$$

A suitable exhaustive summary for the model can be generated containing all the probabilities of the distinct life-histories as separate exhaustive summary terms. This exhaustive summary can then be used to obtain intrinsic parameter redundancy results in Section 4.9 and extrinsic parameter redundancy results in Section 4.10.

Table 4.8: Intrinsic parameter redundancies for capture-recapture-recovery fidelity models

Model	Model rank	Deficiency	Confounded parameters
$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$	4	0	
$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(t)$	n_2+4	0	
$S(\cdot)\ F(\cdot)\ p(t)\ \lambda(\cdot)$	n_2+3	0	
$S(\cdot)\ F(\cdot)\ p(t)\ \lambda(t)$	$2n_2+3$	0	
$S(\cdot)\ F(t)\ p(\cdot)\ \lambda(\cdot)$	n_2+3	0	
$S(\cdot)\ F(t)\ p(\cdot)\ \lambda(t)$	$2n_2+3$	0	
$S(\cdot)\ F(t)\ p(t)\ \lambda(\cdot)$	$2n_2+1$	1	$F_{n_2}p_{n_2+1}$
$S(\cdot)\ F(t)\ p(t)\ \lambda(t)$	$3n_2+1$	1	$F_{n_2}p_{n_2+1}$
$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$	n_2+4	0	
$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(t)$	$2n_2+3$	1	$(1-S_{n_2+1})\lambda_{n_2+1}$
$S(t)\ F(\cdot)\ p(t)\ \lambda(\cdot)$	$2n_2+3$	0	
$S(t)\ F(\cdot)\ p(t)\ \lambda(t)$	$3n_2+2$	1	$(1-S_{n_2+1})\lambda_{n_2+1}$
$S(t)\ F(t)\ p(\cdot)\ \lambda(\cdot)$	$2n_2+3$	0	
$S(t)\ F(t)\ p(\cdot)\ \lambda(t)$	$3n_2+2$	1	$(1-S_{n_2+1})\lambda_{n_2+1}$
$S(t)\ F(t)\ p(t)\ \lambda(\cdot)$	$3n_2+1$	1	$F_{n_2}p_{n_2+1}$
$S(t)\ F(t)\ p(t)\ \lambda(t)$	$4n_2$	2	$F_{n_2}p_{n_2+1}$, $(1-S_{n_2+1})\lambda_{n_2+1}$

4.9 Intrinsic parameter redundancy for the fidelity model

We firstly consider the intrinsic parameter redundancy of the fidelity model in general. Table 4.8 gives us these intrinsic parameter redundancy results for every possibility of parameter dependency in the fidelity model. We denote the fidelity models to be '$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$', where $S(\cdot)$ denotes survival probabilities, $F(\cdot)$ denotes fidelity probabilities, $p(\cdot)$ denotes recapture probabilities and $\lambda(\cdot)$ denotes recovery probabilities, where each different parameter has the option to be either constant, $(\cdot)$, or time-dependent, (t). The second column of Table 4.8 refers to the rank of the model, which is the number of estimable parameters in the model, where there are n_2 recapture occasions in the study with $n_2 \geq 3$. The third column refers to the parameter deficiency of the model where it is parameter redundant when $d > 0$. The final column then gives the confounded parameters if the model is parameter redundant, which are obtained by using the method of solving partial differential equations as detailed in Section 2.2.

The proof of these results can be found in Appendix B.3. Table 4.8 shows that most fidelity models are intrinsically full rank. Furthermore, for the models that are parameter redundant, only the last two sets of parameters are confounded. This is analogous to the result found in Section 4.5 where only the last sets of parameters are

confounded. This intrinsic result for the fully time-dependent model is a confirmation of the result given in Burnham (1993, pp.207) and Cooch and White (2014, Section 9.7) where they indicate there are identifiability problems in the model with the same confounded estimable parameter combinations as given in Table 4.8.

4.10 Extrinsic parameter redundancy for the fidelity model

We now demonstrate how the extrinsic parameter redundancy of a particular data set can be found for the fidelity model using the `Maple` procedure `caprecaprecovfidelity` in Example 4.14 below. We then use this procedure to analyse a study of mountain chicken frogs in Montserrat in Example 4.15.

Example 4.14 - Extrinsic parameter redundancy in the capture-recapture-recovery fidelity model: [See electronic appendix `example4.14.mw`] For the purposes of our model analysis, we require the model to have time-dependency and group-dependency as parameter alternatives. Group-dependency explores how the life-histories differ between groups of animals by allowing parameters to vary among the groups of the animals. For example, if the survival parameters were time- and group-dependent, then the survival probabilities for group '1' would be equal to $S_{1,1}$ for the first year, $S_{1,2}$ for the second year, $S_{1,3}$ for the third year, and so on. The second group would then have parameters $S_{2,1}$, $S_{2,2}$, etc. The procedure `caprecaprecovfidelity` has the parameter options as being constant, time-dependent, group-dependent, or time- and group-dependent for all the parameters in the model. The group which a particular life-history is from is given in the vector **Cov**. Entry i in the vector **Cov** indicates which group the life-history in row i of the data matrix **Data** is from. We see how this group-dependency can be shown in the `Maple` code below.

Consider this data set of distinct life-histories where (i) denotes which group they are from: 1111(1), 1120(2), 1130(3), 1013(1), 1013(3), 1002(2), 1000(1), 1000(2), 1000(3), 0103(1), 0103(2), 0102(1), 0111(1), 0100(1), 0100(2), 0100(3), 0010(1), 0010(2), 0010(3) and 0030(3).

These 20 life-histories are given in the matrix

$$\mathbf{Data} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 0 & 1 & 3 \\ 1 & 0 & 1 & 3 \\ 1 & 0 & 0 & 2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix},$$

with the vector **Cov** indicating which group each respective life-history is from as

$$\mathbf{Cov} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 3 \\ 2 \\ \vdots \\ 2 \\ 3 \\ 3 \end{bmatrix}.$$

Consider a model where the survival parameters are group- and time-dependent, the fidelity parameters are constant, the recapture parameters are time-dependent, and the recovery parameters are group-dependent. We can find the model's exhaustive summary by using the `Maple` code over.

```
kappa := caprecaprecovfidelity2(Data,Cov,4,1,2,3);
# Inputs: (Data,Cov,S,F,z1,z2);
# Data = Data of all life-histories;
# Cov = column vector of which group the life-history from the
#  respective row of Data is from; S = survival probability;
# F = fidelity probability; z1 = recapture probability;
# z2 = recovery probability; for S, F, z1 and z2: '1'=constant,
# '2'=time-dep, '3'=group-dep, '4'=time+group-dep.
```

This code converts the data in the matrix **Data** to generate the model's exhaustive summary as

$$\boldsymbol{\kappa} = \begin{bmatrix} S_{1,1}F^3p_2S_{1,2}p_3S_{1,3}p_4\left\{(1-S_{1,4})(1-\lambda_1)+S_{1,4}\right\} \\ S_{2,1}Fp_2S_{2,2}\left\{(1-F)+F(1-p_3)\right\}(1-S_{2,3})\lambda_2 \\ S_{3,1}F^2p_2S_{3,2}p_3(1-S_{3,3})\lambda_3 \\ S_{1,1}F^3(1-p_2)S_{1,2}p_3S_{1,3}p_4(1-S_{1,4})\lambda_1 \\ S_{3,1}F^3(1-p_2)S_{3,2}p_3S_{3,3}p_4(1-S_{3,4})\lambda_3 \\ S_{2,1}S_{2,2}S_{2,3}\left((1-F)+F(1-p_2)\left[(1-F)\cdots\right.\right. \\ \left.\left.\cdots+\ F(1-p_3)\left\{(1-F)+F(1-p_4)\right\}\right]\right)(1-S_{2,4})\lambda_2 \\ \vdots \\ (1-S_{2,3})(1-\lambda_2)+S_{2,3}(1-F)\left\{(1-S_{2,4})(1-\lambda_2)+S_{2,4}\right\}\cdots \\ \cdots+S_{2,3}F(1-p_4)\left\{(1-S_{2,4})(1-\lambda_2)+S_{2,4}\right\} \\ (1-S_{3,3})(1-\lambda_3)+S_{3,3}(1-F)\left\{(1-S_{3,4})(1-\lambda_3)+S_{3,4}\right\}\cdots \\ \cdots+S_{3,3}F(1-p_4)\left\{(1-S_{3,4})(1-\lambda_3)+S_{3,4}\right\} \\ (1-S_{3,3})\lambda_3 \end{bmatrix}.$$

The parameter deficiency of this exhaustive summary can then be found by the usual methods as described in Section 2.2.

```
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

The parameter set and derivative matrix are given in the `Maple` file `example4.14.mw` only to save space. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to the computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{19}, \mathbf{d} := \mathbf{0}$.

We then conclude that this exemplar data set has no parameter redundancy for the fidelity model with group- and time-dependent survival parameters, constant fidelity parameters, time-dependent recapture parameters, and group-dependent recovery parameters. This means that using that model every parameter in the model can in theory be estimated for that data set. □

Example 4.15 - The mountain chicken frogs data set for the capture-recapture-recovery fidelity model: We now demonstrate extrinsic parameter redundancy in the fidelity model by examining a real data set on mountain chicken frogs in Montserrat. The data were observed during a trial of an anti-fungal drug and the experimenters wanted to analyse survival probabilities between different groups of animals. This study took place for 24 weeks with 15 weeks of treatment followed by 9 weeks of post-treatment monitoring, however a volcanic eruption cut the study short at that point, when the experimenters were intending to continue the study further. All of the individual animals were split into three different groups: A control group of frogs who were caught and checked to not have the disease, a control group of frogs who were bathed in a stream water bath to replicate the stresses of the treatment but without the drug, and a treatment group of frogs which were treated with the drug. We denote these as group '1' to be a control group, group '2' to be a bath-control group, and group '3' to be the group which has had the anti-fungal treatment. Note that this study is not a randomised experiment so the amount of possible life-histories for each group is dependent on how many frogs were selected for each group by the experimenters. There are 36 distinct life-histories from the control group, 87 distinct life-histories from the bath-control group and 67 distinct life-histories from the treatment group. The experimenters had two aims for this study: First they wanted to see if there was a difference between the survival probabilities before and after the treatment of the drug after 15 weeks, and they secondly wanted to see if the animals treated with the drug had higher survival probabilities than a control group of animals. We use this data set as an illustration of extrinsic parameter redundancy results for the fidelity model even though it is unclear if the experimenters had considered immigration to be a factor in this study.

We can use the previous `Maple` procedure `caprecaprecovfidelity` to produce an exhaustive summary so that extrinsic parameter redundancy results can be obtained for the mountain chicken frogs data for all of the different possible fidelity models. We give these results in Table 4.9, where we denote the models as '$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$', where $S(\cdot)$ denotes survival probabilities, $F(\cdot)$ denotes fidelity probabilities, $p(\cdot)$ denotes re-

capture probabilities and $\lambda(\cdot)$ denotes recovery probabilities. Each different parameter has the options to be either constant, $(\cdot)$, time-dependent, (t), or group-dependent, (g), e.g. $S(g)$ indicates a model where the survival parameters are group-dependent. We do not list any results exploring when parameters are both group- and time-dependent, as all of these models are parameter redundant with a parameter deficiency at least of 5.

As we can see in Table 4.9, every fidelity model that is intrinsically full rank still remains full rank when exploring the mountain chicken frogs data set. These results show that many fidelity models are not parameter redundant even when examining sparse data sets. However we exclude the cases where a set of parameters are time- and group-dependent, as when this occurs the data are not rich enough so that the model is then parameter redundant. □

4.11 Discussion

We have advanced our theory on capture-recapture models from Chapter 3 to now include dead recoveries in the same study. In this case an animal can be recaptured many times before being found dead, giving us estimates of recapture, recovery and survival probabilities. Using joint capture-recapture-recovery models is better than considering two separate mark-recovery and capture-recapture analyses as the survival estimate are more precise (Burnham, 1993). We considered intrinsic parameter redundancy results of these models with different time- and age- dependencies in Section 4.5 by first showing simper exhaustive summaries in Theorem 3.2 which are proved in Appendix B.2. Many of the models are intrinsically not parameter redundant and even for the models that are parameter redundant, these models only confound the last set of parameters for the survival, recapture and recovery parameters for the final recapture/recovery occasion n_2 of the study. This is advantageous as while a model may still be parameter redundant, the majority of the survival parameters can be estimated without the need for constraining parameters or reparameterisation.

Table 4.9: Extrinsic parameter redundancies for capture-recapture-recovery fidelity models for the mountain chicken frogs data set

Model	Rank	Deficiency	Model	Rank	Deficiency	Model	Rank	Deficiency
$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$	4	0	$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(t)$	27	0	$S(\cdot)\ F(\cdot)\ p(\cdot)\ \lambda(g)$	6	0
$S(\cdot)\ F(\cdot)\ p(t)\ \lambda(\cdot)$	26	0	$S(\cdot)\ F(\cdot)\ p(t)\ \lambda(t)$	49	0	$S(\cdot)\ F(\cdot)\ p(t)\ \lambda(g)$	28	0
$S(\cdot)\ F(\cdot)\ p(g)\ \lambda(\cdot)$	6	0	$S(\cdot)\ F(\cdot)\ p(g)\ \lambda(t)$	29	0	$S(\cdot)\ F(\cdot)\ p(g)\ \lambda(g)$	8	0
$S(\cdot)\ F(t)\ p(\cdot)\ \lambda(\cdot)$	26	0	$S(\cdot)\ F(t)\ p(\cdot)\ \lambda(t)$	49	0	$S(\cdot)\ F(t)\ p(\cdot)\ \lambda(g)$	28	0
$S(\cdot)\ F(t)\ p(t)\ \lambda(\cdot)^*$	47	1	$S(\cdot)\ F(t)\ p(t)\ \lambda(t)^*$	70	1	$S(\cdot)\ F(t)\ p(t)\ \lambda(g)^*$	49	1
$S(\cdot)\ F(t)\ p(g)\ \lambda(\cdot)$	28	0	$S(\cdot)\ F(t)\ p(g)\ \lambda(t)$	51	0	$S(\cdot)\ F(t)\ p(g)\ \lambda(g)$	30	0
$S(\cdot)\ F(g)\ p(\cdot)\ \lambda(\cdot)$	6	0	$S(\cdot)\ F(g)\ p(\cdot)\ \lambda(t)$	29	0	$S(\cdot)\ F(g)\ p(\cdot)\ \lambda(g)$	8	0
$S(\cdot)\ F(g)\ p(t)\ \lambda(\cdot)$	28	0	$S(\cdot)\ F(g)\ p(t)\ \lambda(t)$	51	0	$S(\cdot)\ F(g)\ p(t)\ \lambda(g)$	30	0
$S(\cdot)\ F(g)\ p(g)\ \lambda(\cdot)$	8	0	$S(\cdot)\ F(g)\ p(g)\ \lambda(t)$	31	0	$S(\cdot)\ F(g)\ p(g)\ \lambda(g)$	10	0
$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$	27	0	$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(t)^*$	49	1	$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(g)$	29	0
$S(t)\ F(\cdot)\ p(t)\ \lambda(\cdot)$	49	0	$S(t)\ F(\cdot)\ p(t)\ \lambda(t)^*$	71	1	$S(t)\ F(\cdot)\ p(t)\ \lambda(g)$	51	0
$S(t)\ F(\cdot)\ p(g)\ \lambda(\cdot)$	29	0	$S(t)\ F(\cdot)\ p(g)\ \lambda(t)^*$	51	1	$S(t)\ F(\cdot)\ p(g)\ \lambda(g)$	31	0
$S(t)\ F(t)\ p(\cdot)\ \lambda(\cdot)$	49	0	$S(t)\ F(t)\ p(\cdot)\ \lambda(t)^*$	57	15	$S(t)\ F(t)\ p(\cdot)\ \lambda(g)$	51	0
$S(t)\ F(t)\ p(t)\ \lambda(\cdot)^*$	70	1	$S(t)\ F(t)\ p(t)\ \lambda(t)^*$	78	16	$S(t)\ F(t)\ p(t)\ \lambda(g)^*$	72	1
$S(t)\ F(t)\ p(g)\ \lambda(\cdot)$	51	0	$S(t)\ F(t)\ p(g)\ \lambda(t)^*$	60	14	$S(t)\ F(t)\ p(g)\ \lambda(g)$	53	0
$S(t)\ F(g)\ p(\cdot)\ \lambda(\cdot)$	29	0	$S(t)\ F(g)\ p(\cdot)\ \lambda(t)^*$	51	1	$S(t)\ F(g)\ p(\cdot)\ \lambda(g)$	31	0
$S(t)\ F(g)\ p(t)\ \lambda(\cdot)$	51	0	$S(t)\ F(g)\ p(t)\ \lambda(t)^*$	73	1	$S(t)\ F(g)\ p(t)\ \lambda(g)$	53	0
$S(t)\ F(g)\ p(g)\ \lambda(\cdot)$	31	0	$S(t)\ F(g)\ p(g)\ \lambda(t)^*$	53	1	$S(t)\ F(g)\ p(g)\ \lambda(g)$	33	0
$S(g)\ F(\cdot)\ p(\cdot)\ \lambda(\cdot)$	6	0	$S(g)\ F(\cdot)\ p(\cdot)\ \lambda(t)$	29	0	$S(t)\ F(\cdot)\ p(\cdot)\ \lambda(g)$	8	0
$S(g)\ F(\cdot)\ p(t)\ \lambda(\cdot)$	28	0	$S(g)\ F(\cdot)\ p(t)\ \lambda(t)$	51	0	$S(t)\ F(\cdot)\ p(t)\ \lambda(g)$	30	0
$S(g)\ F(\cdot)\ p(g)\ \lambda(\cdot)$	8	0	$S(g)\ F(\cdot)\ p(g)\ \lambda(t)$	31	0	$S(t)\ F(\cdot)\ p(g)\ \lambda(g)$	10	0
$S(g)\ F(t)\ p(\cdot)\ \lambda(\cdot)$	28	0	$S(g)\ F(t)\ p(\cdot)\ \lambda(t)$	51	0	$S(t)\ F(t)\ p(\cdot)\ \lambda(g)$	30	0
$S(g)\ F(t)\ p(t)\ \lambda(\cdot)^*$	49	1	$S(g)\ F(t)\ p(t)\ \lambda(t)^*$	72	1	$S(t)\ F(t)\ p(t)\ \lambda(g)^*$	51	1
$S(g)\ F(t)\ p(g)\ \lambda(\cdot)$	30	0	$S(g)\ F(t)\ p(g)\ \lambda(t)$	53	0	$S(t)\ F(t)\ p(g)\ \lambda(g)$	32	0
$S(g)\ F(g)\ p(\cdot)\ \lambda(\cdot)$	8	0	$S(g)\ F(g)\ p(\cdot)\ \lambda(t)$	31	0	$S(t)\ F(g)\ p(\cdot)\ \lambda(g)$	10	0
$S(g)\ F(g)\ p(t)\ \lambda(\cdot)$	30	0	$S(g)\ F(g)\ p(t)\ \lambda(t)$	53	0	$S(t)\ F(g)\ p(t)\ \lambda(g)$	32	0
$S(g)\ F(g)\ p(g)\ \lambda(\cdot)$	10	0	$S(g)\ F(g)\ p(g)\ \lambda(t)$	33	0	$S(t)\ F(g)\ p(g)\ \lambda(g)$	12	0
All the fidelity models that are starred* are intrinsically parameter redundant								

We then considered extrinsic parameter redundancy results by exploring a data set of cormorants from Hènaux et al. (2007). We consider three subsets of the data which vary in sparseness and observe how this affects the ability to estimate parameters. Even for the relatively sparse data set of the colony 1 data only where only 465 out of the possible 12,282 distinct life-histories were recorded, there are very few models that are parameter redundant. This shows that even a sparse data set may have enough data to be able to estimate all the parameters in a capture-recapture-recovery model. We also consider generalising the life-histories in Section 4.7 by considering a sparseness statistic c so that only the life-histories that have c or fewer occasions between first capture and last recapture/recovery are considered. This extrinsic case is the more likely to occur in practice as very rarely will there be a perfect data set observed.

We further went on to explore a different capture-recapture-recovery model in Sections 4.8 to 4.10 based on Burnham (1993, Section 3.2). From this fidelity model we can obtain estimates of the probability of animal emigration as well as the probability of animal survival. We explored how the probabilities of these life-histories can be expressed and gave some intrinsic parameter redundancy results as well as examples of extrinsic parameter redundancies including exploring a data set on mountain chicken frogs in Montserrat.

Finally, it is worth noting that some of the models in this chapter have some assumptions which need to be satisfied to achieve unbiased parameter estimates. Many of these assumptions are similar to those mentioned in Sections 2.1 and 3.9.

1. Each animal has an equal recapture and recovery probability. This may not be the case due to animals being trap-happy or trap-shy, which may give a negative or positive estimate bias.

2. We do not consider the loss of marking tags during the study. Potentially some of animals may lose their tags during the study, and this may influence the number of recaptures/recoveries recorded. Tag loss can be incorporated into a capture-recapture-recovery model by the use of a parameter (such as shown in Arnason and Mills, 1981, and Pomeroy et al., 2010) but such models have not been considered in parameter redundancy analysis in a capture-recapture-recovery framework.

3. The discrete time period between recovery occasions needs to be same during the study. As we are working on a discrete time period this generally means yearly

intervals, and these time periods have to remain at yearly intervals throughout the study. Capture-recapture-recovery models are being developed for the continuous time case, such as in Langrock and King (2013).

4. Technological and/or human errors in the data are not considered, or at least are considered to be of negligible value. This includes issues like incorrect individual identification and incorrect data input into computers. Note that this is different from having missing observations as we are saying here that there are no *incorrect* observations rather than missing ones.

5. We assume that the recapturing and recovering of animals does not have an effect on the survival probabilities of other animals in the study. Similarly, the recapturing and recovering of animals does not have an effect on the probability of recapture or recovery for the other animals in the study as well.

We now leave models where we mark or capture individuals, for a different class of models called occupancy models, which are explored in Chapter 5.

Chapter 5

Occupancy Models

5.1 Occupancy modelling background

The previous chapters considered models involving marking of individual animals and then recapturing and/or recovering them later on in the study. Rather than individuals being marked, this chapter examines models for data where the presence and absence of a species is recorded, which is known as **occupancy modelling**. These occupancy studies can be used to estimate the proportion of sites which are occupied by a certain species. Collections of occupancy data can form a basis of a monitoring program where the probability of occupancy at certain sites is closely supervised. As MacKenzie et al. (2006) state, we, "did not appreciate the generality of the concept of 'occupancy'," until the development of these occupancy models in the last decade. Before the formulation of these occupancy models in the early 2000's, it was considered difficult to estimate absolute abundance change among large sites as marking individual animals proved problematic. The parameters in occupancy models are different from the ones in Chapters 2, 3 and 4, as instead of estimating the probability of survival, the occupancy model estimates the probability that a species is present at a location. Estimating species occupancy can be advantageous in some ecological studies where it is easier to collect data on whether a species is present or absent compared to the effort to mark and recapture/recover individuals.

The first occupancy models in animal studies used ad-hoc methods to estimate occupancy probabilities where the detection probabilities were less than one, such as in Geissler and Fuller (1987) and Azuma et al. (1990). The beginnings for the increased use of occupancy studies comes from the work of MacKenzie et al. (2002, 2003) and Tyre et al. (2003). MacKenzie et al. (2002) and Tyre et al. (2003) explore a model with

replicated sampling surveys where observers would see if they can detect a species during each survey. A survey is an attempt by an observer to detect an animal species at a given time point. The replication of these surveys can form the basis of a monitoring program which provides information about the presence or absence of animal species at that time. The model likelihoods of MacKenzie et al. (2002) and Tyre et al. (2003) account for imperfect detection by noting that the non-detection of a species does not necessarily imply that the species is absent. MacKenzie et al. (2003) extends this work to consider a model with multiple seasons, where each season would be defined as a different time period during which the presence or absence of a species is assumed to be static. MacKenzie et al. (2003) consider changes in occupancy probability between seasons, such as the species going from being present to absent, or vice-versa. MacKenzie et al. (2003) has been widely referenced since as it describes the most employed occupancy model at the current time due to its simplicity and flexibility of use. The book by MacKenzie et al. (2006) has also been widely cited due to the wide range of models discussed in their book, and a review of the recent advances in occupancy modelling can be seen in Bailey et al. (2014), which was presented at the 2013 EURING meeting in Georgia, USA.

Recent applications of occupancy methodology have been wide ranging in their use such as in Bailey et al. (2004a) monitoring terrestrial salamanders, in O'Connell Jr. et al. (2006) using the MacKenzie et al. (2002) model to derive parameter estimates of occupancy for a range of mammals in Cape Cod, Massachusetts, and in Falke et al. (2010) which explores Great Plains stream fish in North America using the multiple state occupancy model from Nichols et al. (2007). The area of species distribution modelling, in which presence/absence studies play a role, is also a recently developing field which considers results of occupancy studies to predict such aspects as conservation areas in a country. The application of these methods is shown in Beale et al. (2012), or where potentially unobserved species have not been found yet, as in Pearson et al. (2007). It is possible to combine different data sets to improve the accuracy of occupancy estimates as well, as shown in Nichols et al. (2008). There has also been a wide range of extensions to the original occupancy models of MacKenzie et al. (2002, 2003), such as observing species in multiple occupancy states in MacKenzie et al. (2009), observing multiple interacting species in MacKenzie et al. (2004), and the use of occupancy models to derive estimates of animal abundance in Royle and Nichols (2003). We explore parameter redundancy results for these three occupancy model extensions in Sections 6.1, 6.3 and 6.4 respectively later on in this thesis.

Occupancy models can be fitted using `PRESENCE` (MacKenzie and Hines, 2014) which is a software package developed to derive parameter estimates for a range of occupancy models, though most occupancy models can also be fitted using the software package `MARK` (Cooch and White, 2014). Alternative software programs that also fit occupancy models include `E-SURGE` (Gimenez et al., 2014), and `R` using the package `unmarked` (Fiske and Chandler, 2011, and Fiske et al., 2014).

We start by considering parameter redundancy in the basic occupancy model in Section 5.2 which has only one survey in only one season. Then this can be extended to include multiple detection surveys in Section 5.3 and multiple seasons in Section 5.4, as well as observing species from multiple groups in Section 5.7. Intrinsic parameter redundancy results are obtained in Section 5.5, and extrinsic parameter redundancy is explored in Sections 5.6 and 5.8. There is also a final section on the effect of data sparseness in Section 5.9.

5.2 The basic occupancy model

The simplest occupancy model is the model which has only one survey which detects whether the species is present or absent. We consider adding more surveys in the next section but consider only one survey for now. We let p denote the probability of species detection at that survey. Observe that p denotes a different probability from the previous sections as p now indicates when the species is detected, and *not* when an individual is recaptured in a capture-recapture or capture-recapture-recovery model. The other parameter required in this model is the probability of species occupation, denoted as φ. This species occupation parameter gives the probability that the species is present at the beginning of the study.

For this simplest occupancy model, it is clear there are only two possibilities: Either the species is detected, or the species is not detected. If the species is detected then it must be present at this site and has the probability φp. If the species is not detected however, it is not clear if the species is absent at that site, or if the species is actually present but was not detected. This problem of non-detection is important to consider as a non-detection does not necessarily imply the species is absent at that site. If we do not consider the case where the species could be present but is not detected, we achieve bias in our estimates, especially in the underestimation of the true occupancy rates (MacKenzie et al., 2006, pp.104-105). The probability of the species being undetected is the probability the species was present but not detected, $\varphi(1-p)$, plus the probability

the species was absent, $(1-\varphi)$. An exhaustive summary then includes the probabilities of these two options, $\boldsymbol{\kappa} = [\varphi p, \varphi(1-p) + (1-\varphi)]^T$. This model however is parameter redundant, as when we find the derivative matrix with respect to its parameters $\boldsymbol{\theta} = [\varphi, p]^T$ we obtain that

$$\mathbf{D} = \left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} \varphi & -\varphi \\ p & -p \end{bmatrix},$$

only has a matrix rank of 1, meaning the exhaustive summary has a parameter deficiency of 1. This is clear to see as we observe that $\varphi(1-p) + (1-\varphi) = 1 - \varphi p$ so that the parameters p and φ only ever appear as the product φp. This is well known in MacKenzie et al. (2003) and the idea behind the robust design is that when more than one survey is conducted, this parameter redundancy is removed. We show this in the next section.

5.3 Adding more detection surveys

We can extend the previous simple model by considering when observers attempt to detect animal species at more than one survey during the study, see for example MacKenzie et al. (2006, Section 4.4). We are assuming that the occupancy status for a species does not change between surveys, so that a species that is present remains present during the whole of the study, and similarly for a species which is absent.

In this model we still denote the probability of species occupancy as φ, but we now denote the probability of detection as p_j for survey $j = 1, \ldots, k$, where k is the total number of surveys in the study. We follow a similar notation to capture-recapture studies and denote a particular series of detections in an **occupancy-history** as a binary series of 1's and 0's. We let '1' denote when the species was detected for that particular survey and '0' denote when the species was *not* detected. Examples 5.1 and 5.2 below show how the probability of particular occupancy-histories can be obtained.

Example 5.1: A possible occupancy-history could be $h = 01110$. This means that the species was not detected during surveys 1 and 5 in the study, but was detected during surveys 2, 3 and 4. Observe that the occupancy-history does not need a detection to begin, as unlike the models of Chapter 2, 3 and 4, we do not need to mark the individuals first before detection. As the species was detected in at least one survey it shows the species is present at that site, and the probability of this occupancy-history is then equal to

$$Pr(h) = \varphi(1-p_1)p_2p_3p_4(1-p_5).$$

□

Example 5.2: The probability of the occupancy-history where there are no detections in the study denoted as $h = 00000$, is equal to

$$Pr(h) = \varphi\left(1-p_1\right)\left(1-p_2\right)\left(1-p_3\right)\left(1-p_4\right)\left(1-p_5\right) + \left(1-\varphi\right).$$

This is due to considering both cases where the species could be present but not detected at all during the study, or whether the species is absent. □

The probabilities of all the observed occupancy-histories forms an exhaustive summary for the occupancy model. If we let δ_j represent the individual occupancy-history entry for survey j (which is either equal to zero or one for all j), then the probability of occupancy-history h_X occurring is

$$\Pr(h_X) = \begin{cases} \varphi \prod_{j=1}^{k} \left\{\delta_j p_j + \left(1-\delta_j\right)\left(1-p_j\right)\right\} & \text{if } \sum_{j=1}^{k} \delta_j \geq 1, \\ \varphi \prod_{j=1}^{k} \left(1-p_j\right) + \left(1-\varphi\right) & \text{if } \sum_{j=1}^{k} \delta_j = 0. \end{cases} \tag{5.1}$$

All the possible occupancy-histories fall into the top expression in Equation (5.1) apart from the history which has no detections at all during the study. The exhaustive summary is then a vector of all the occupancy-history probabilities recorded in the study. This occupancy model can also be referred to as being a zero-inflated binomial model, as stated in Royle and Nichols (2003), as occupancy data commonly contains a large of non-detections during a study. Zero-inflated models have been used before in statistical ecology such as in Ridout et al. (2001) and Morgan (2008, pp.124).

If we consider the intrinsic parameter redundancy of this occupancy model, this model is *not* parameter redundant and all of the parameters in the model are estimable,

given that at least two surveys are performed. When there is only one survey, this is the same as the basic simplest model of Section 5.2 which was shown to be parameter redundant with a deficiency of one. This is a result which is proved later on in Theorem 5.1 in Section 5.5.

However, typically the use of occupancy modelling is to examine presence/absence patterns over multiple time periods. Therefore, we consider the parameter redundancy of occupancy studies performed over a multiple number of seasons, and this model will be described in the next section.

5.4 Increasing the number of seasons

Rather than one season with multiple surveys, a typical study will observe a site over a multiple number of seasons. This **multiple-seasons model** is more widely used in practice due to the estimation of parameters where the species possibly change occupancy states between seasons. This change of occupancy status from present to absent, or vice-versa, may be of interest when considering if a species occupies the site over particular time periods, and is the reason why this model is regularly used in occupancy studies.

Let T represent the number of seasons observed in the study at a single site, with k_t as the number of surveys during season t. k_t is generally the same number of surveys for every season in the study but this does not necessarily have to be the case. As this is a two state model where the only states are presence and absence, we define state '1' to be where the species is present and state '0' to be where the species is absent. Let $p_{t,j}$ denote the probability of detection during survey j during season t for $j = 1, \ldots, k_t$ and $t = 1, \ldots, T$. Let $\varphi_0^{[1]}$ be the probability that the species is present at time $t = 0$, and $\varphi_t^{[a,1]}$ be the transition probability that a site that is occupancy state a in season t, where $a = \{0, 1\}$, is then occupied by the target species in season $t + 1$ for all $t = 1, \ldots, T - 1$.

Note that MacKenzie et al. (2003) uses a different set of parameters to define these transition probabilities. They define $\epsilon_t = 1 - \varphi_t^{[1,1]}$ as the probability of moving from being present in season t to being absent in season $t + 1$ (i.e. moving from state 1 to state 0), also known as species *extinction*. They further define as $\gamma_t = \varphi_t^{[0,1]}$ as the probability of moving from being absent in season t to being present in season $t + 1$ (i.e. moving from state 0 to state 1), also known as species *colonisation*. Due to the

reparameterisation theorem Theorem 2.3, this reparameterisation results in the same parameter redundancy as the parameterisation in terms of $\varphi_t^{[1,1]}$ and $\varphi_t^{[0,1]}$. Here we use the $\varphi_t^{[a,1]}$ notation as the model more naturally extends to the cases with more states as explored later in Section 6.1. It is also worth noting that $\varphi_0^{[0]} = 1 - \varphi_0^{[1]}$, and $\varphi_t^{[a,0]} = 1 - \varphi_t^{[a,0]}$ for each state $a = \{0, 1\}$, to enable the model to have the least number of parameters possible.

We assume that species extinction or species colonisation does not happen mid-season and only happen between seasons, with each season denoting the same discrete time period throughout the study. We further assume that the dynamic processes between seasons are being explicitly modelled, though it is possible to have an implicitly modelled study by defining $\varphi_t^{[a,1]} = \varphi_t$ for all t, i.e. the state it was in previously is not considered. Such a model is shown in MacKenzie et al. (2006, Section 7.2), however we explore explicit models only in this thesis where the previous state in the last season is considered.

Let ϕ_t denote the transition matrix

$$\phi_t = \begin{bmatrix} \varphi_t^{[1,1]} & 1 - \varphi_t^{[1,1]} \\ \varphi_t^{[0,1]} & 1 - \varphi_t^{[0,1]} \end{bmatrix},$$

for all $t = 1, \ldots, T-1$, and let ϕ_0 be the row vector corresponding to the initial species presence/absence as

$$\phi_0 = \begin{bmatrix} \varphi_0^{[1]} & 1 - \varphi_0^{[1]} \end{bmatrix}.$$

Let $\delta_{t,j}$ denote the species observed state during season t in survey j, where $\delta_{t,j} = 1$ when the species is detected and $\delta_{t,j} = 0$ when the species is not detected. Let $\mathbf{p}_{X,t}$ be the column vector where each entry is the probability of observing the occupancy-history h_X in season t conditional upon its occupancy state. The general form for $\mathbf{p}_{X,t}$ can be expressed as

$$\mathbf{p}_{X,t} = \begin{bmatrix} \prod_{j=1}^{k_t} \left\{ \delta_{t,j} p_{t,j} + (1 - \delta_{t,j}) (1 - p_{t,j}) \right\} \\ \prod_{j=1}^{k_t} (1 - \delta_{t,j}) \end{bmatrix}. \tag{5.2}$$

This does mean that if there is at least one detection during season t, the second entry in $\mathbf{p}_{X,t}$ would be equal to zero. This shows that is impossible for a species to be absent

from the site if the species has been detected during the season. We show examples of $\mathbf{p}_{X,t}$ vectors in Examples 5.1 and 5.2 below.

Example 5.1 revisited: The $\mathbf{p}_{X,t}$ vector with the occupancy-history $h = 01110$ in season t would be equal to

$$\mathbf{p}_{01110,t} = \begin{bmatrix} (1-p_{t,1})p_{t,2}p_{t,3}p_{t,4}(1-p_{t,5}) \\ 0 \end{bmatrix}.$$

□

Example 5.2 revisited: The $\mathbf{p}_{X,t}$ vector with the occupancy-history $h = 00000$ which has no detections in the season t would be equal to

$$\mathbf{p}_{00000,t} = \begin{bmatrix} (1-p_{t,1})(1-p_{t,2})(1-p_{t,3})(1-p_{t,4})(1-p_{t,5}) \\ 1 \end{bmatrix}.$$

□

Let $\Delta(\mathbf{p}_{X,t})$ denote a diagonal matrix with diagonal entries given by the vector $\mathbf{p}_{X,t}$ and zeros on the non-diagonal entries. Note that this is the same matrix as denoted by $D(\mathbf{p}_{X,t})$ from in MacKenzie et al. (2003) so that we do not confuse this diagonal matrix Δ with our derivative matrix $\mathbf{D}$.

Example 5.2 revisited: The matrix $\Delta(\mathbf{p}_{00000,t})$ for $t \neq T$ is given as

$$\Delta(\mathbf{p}_{0000,t}) = \begin{bmatrix} (1-p_{t,1})(1-p_{t,2})(1-p_{t,3})(1-p_{t,4})(1-p_{t,4}) & 0 \\ 0 & 1 \end{bmatrix}.$$

□

We observe that no diagonal matrix is required for the detections in the final season T in the study to give the correct matrix dimensions. The probability of the occupancy-history h_X is given as

$$Pr(h_X) = \phi_0 \left[\prod_{t=1}^{T-1} \Delta(\mathbf{p}_{X,t})\phi_t \right] \mathbf{p}_{X,T}. \tag{5.3}$$

The likelihood for the multiple-seasons occupancy model is then

$$L = \prod_{X=1}^{N} Pr(h_X), \tag{5.4}$$

for the N different occupancy-histories observed in a study. When more seasons are added in a model, we will use a vertical dash '|' to indicate the change in seasons in the binary code for the occupancy-history, though some papers use blanks in the code to indicate this, as in MacKenzie et al. (2003, 2009). We give Example 5.3 and 5.4 below to show how the probability of an occupancy-history can be obtained.

Example 5.3: Consider a study where we have 3 seasons where there are 2 surveys in every season. The probability of observing the occupancy-history $h = 10|11|01$, where the species is detected once in the first season in the first survey, is detected both times in season two, and is only detected in the last survey in the final season, can be found by multiplying the following matrices and vectors together:

$$\begin{aligned} Pr(h) &= \phi_0\Delta(\mathbf{p}_{10,1})\phi_1\Delta(\mathbf{p}_{11,2})\phi_2\mathbf{p}_{01,3} \\ &= \begin{bmatrix} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}(1-p_{1,2}) & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} p_{2,1}p_{2,2} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1})p_{3,2} \\ 0 \end{bmatrix} \\ &= \varphi_0^{[1]}p_{1,1}(1-p_{1,2})\varphi_1^{[1,1]}p_{2,1}p_{2,2}\varphi_2^{[1,1]}(1-p_{3,1})p_{3,2}. \end{aligned}$$

□

Example 5.4: The probability of the occupancy-history $h = 00|00|00$, where there are no detections at all for any surveys in the study, is

$$\begin{aligned} Pr(h) &= \phi_0\Delta(\mathbf{p}_{00,1})\phi_1\Delta(\mathbf{p}_{00,2})\phi_2\mathbf{p}_{00,3} \\ &= \begin{bmatrix} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{1,1})(1-p_{1,2}) & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} (1-p_{2,1})(1-p_{2,2}) & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1})(1-p_{3,2}) \\ 1 \end{bmatrix} \\ &= \varphi_0^{[1]}(1-p_{1,1})(1-p_{1,2}) \left[\varphi_1^{[1,1]}(1-p_{2,1})(1-p_{2,2}) \left\{\varphi_2^{[1,1]}(1-p_{3,1})(1-p_{3,2}) \right.\right. \\ &+ \left.\left. \left(1-\varphi_2^{[1,1]}\right)\right\} + \left(1-\varphi_1^{[1,1]}\right) \left\{\varphi_2^{[0,1]}(1-p_{3,1})(1-p_{3,2}) + \left(1-\varphi_2^{[0,1]}\right)\right\}\right] \\ &+ \left(1-\varphi_0^{[1]}\right) \left[\varphi_1^{[0,1]}(1-p_{2,1})(1-p_{2,2}) \left\{\varphi_2^{[1,1]}(1-p_{3,1})(1-p_{3,2}) \right.\right. \\ &+ \left.\left. \left(1-\varphi_2^{[1,1]}\right)\right\} + \left(1-\varphi_1^{[0,1]}\right) \left\{\varphi_2^{[0,1]}(1-p_{3,1})(1-p_{3,2}) + \left(1-\varphi_2^{[0,1]}\right)\right\}\right]. \end{aligned}$$

□

The complex expression in Example 5.4 stems from the fact that non-detection in a

season does not necessarily mean the species is absent at that time period and the true occupancy-history behind the species could have 8 different possibilities depending on colonisations and/or extinctions between seasons.

As in the previous chapters in this thesis, we note that this occupancy model has underlying assumptions. A number of these assumptions are discussed in MacKenzie et al. (2002, 2003) and Mazerolle et al. (2007).

1. The occupancy state of the species does not change within a season. While we have parameters to indicate the probability of when a species becomes extinct or is colonised at a site, these parameters are only given between seasons. It is possible a species could become extinct half way through a season, but accounting for this in the model may possibly mean the model becomes parameter redundant. We assume in these models that extinction and colonisation happens between the last survey of one season and the first survey of the next season. Otto et al. (2013) considers this assumption and shows there is bias in the parameter estimates when this assumption does not hold in occupancy studies.

2. Detection and non-detection of species does not have an effect on future probabilities in the model. Detecting (or not detecting) a species at a certain time point should not influence future detection probabilities or future extinction and colonisation probabilities either if there is more than one season in the study. An example where this assumption could be violated is if detecting a species in some way injures or disturbs the detected individual or habitat, which then makes detection of the species less likely in the next survey. Still, if the individual numbers are large enough this could be considered of negligible value.

3. The sampling intervals of each season are all of the same time period during the study. As we are working on a discrete time period, seasons generally mean yearly intervals, and these have to remain yearly intervals throughout the study.

4. Technological and/or human errors in the data collection are considered to be negligible. This includes items like incorrect species identification or incorrect data input into computers. Note that this is somewhat different from having missing observations, as we will discuss in Section 5.8, as we are saying here there are no *incorrect* observations rather than missing ones.

We begin our analysis of occupancy models by generating a result for intrinsic parameter redundancy in the next section.

5.5 Intrinsic parameter redundancy for occupancy models

To examine intrinsic parameter redundancy in the multiple capture-recapture and capture-recapture-recovery models in Sections 3.4 and 4.4, we found that there were simpler exhaustive summaries that could be used. This reduced the large number of exhaustive summary terms in the model considerably. We can then go on to obtain general intrinsic parameter redundancy results using this simpler exhaustive summary. There is a similar result for occupancy models as given in Theorem 5.1 below.

Theorem 5.1. *a) The standard two-state occupancy model with perfect data for any number of surveys per season greater than one, is not parameter redundant. Any season with only one survey or no surveys at all during the season adds one to the parameter deficiency of that model.*
b) A simpler exhaustive summary of the occupancy model consists of the terms

- $s_{t,j} = \varphi_{t-1}^{[1,1]} p_{t,j}$ *for all* $t = 2, \ldots, T$ *and* $j = 1, \ldots, k_t$, *where* $s_{1,j} = \varphi_0^{[1]} p_{1,j}$ *for all* $j = 1, \ldots, k_t$,
- $r_t = \varphi_{t-1}^{[1,1]} \prod_{j=1}^{k_t} p_{t,j}$ *for all* $t = 2, \ldots, T$ *and* $j = 1, \ldots, k_t$, *where* $r_1 = \varphi_0^{[1]} \prod_{j=1}^{k_1} p_{1,j}$ *for all* $j = 1, \ldots, k_t$,
- *and* $u_t = \varphi_t^{[0,1]} / \varphi_t^{[1,1]}$ *for all* $t = 1, \ldots, T-1$.

The proof of Theorem 5.1 can be found in Appendix B.4. The proof uses the two-stage extension theorem of Theorem 3.1 to prove there is a simpler exhaustive summary as given in the theorem, and then we can use the standard extension theorem of Theorem 2.2 to prove the simpler exhaustive summary is full rank for a general number of surveys and seasons. This theorem proves that there is no intrinsic parameter redundancy in the standard two-state occupancy model, given that there are at least two surveys in each season.

We note that the number of exhaustive summary terms grows exponentially with number of surveys and seasons during a study, with rate $2^{T \times \sum_{t=1}^{T} k_t}$. In practice for a large number of seasons, not all of the possible occupancy-histories will be observed in a typical data set. It is for this reason why we consider extrinsic parameter redundancy results for occupancy models in the next section.

5.6 Extrinsic parameter redundancy for occupancy models

Similarly to Sections 3.7, 4.6 and 4.10, we have developed `Maple` code which considers the extrinsic parameter redundancy of a given data set. This `Maple` procedure, `occmodseasonshist`, converts data of observed occupancy-histories into exhaustive summary terms which can be examined for parameter redundancy. This procedure uses matrix multiplication to generate each exhaustive summary term corresponding to its occupancy-history. We can further relax our dependencies on the model parameters as follows:

- By relaxing season-dependence on the between seasons occupancy parameters: $\varphi_t^{[1,1]} = \varphi_c^{[1,1]}$ and $\varphi_t^{[0,1]} = \varphi_c^{[0,1]}$ for all seasons $t = 1, \ldots, T-1$. We give the subscript c to denote a constant occupancy probability over all seasons. Note that $\varphi_0^{[1]}$ remains as while other transition probabilities go from state a to state 1, $\varphi_0^{[1]}$ has no previous state and cannot be seen to be a transition probability. This is viewed in Section 6.2 where we extend the model to include a given previous state at time $t = 0$.

- By relaxing survey-dependence on the detection parameters: $p_{t,j} = p_t$ for all surveys $j = 1, \ldots, k_t$ and seasons $t = 1, \ldots, T$. There is then no difference between detection probabilities within a single season.

- By relaxing season-dependence on the detection parameters: $p_{t,j} = p_j$ for all surveys $j = 1, \ldots, k_t$ and seasons $t = 1, \ldots, T$. There is then no difference between detection probabilities between seasons, and the probabilities are only dependent on which survey is considered during the season.

- By relaxing both season-dependence and survey-dependence on the detection parameters: $p_{t,j} = p$ for all surveys $j = 1, \ldots, k_t$ and seasons $t = 1, \ldots, T$. There is then no difference between any detection probabilities for the whole of the study.

We use the model notation of $\varphi()p()$ to distinguish between the different parameter dependencies in a model. The transition occupancy parameters $\varphi_t^{[a,1]}$ can be either constant, $\varphi(\cdot)$, or season-dependent, $\varphi(t)$. The detection parameters $p_{t,j}$ can be either constant, $p(\cdot)$, season-dependent, $p(t)$, survey-dependent, $p(j)$, or season- and survey-dependent, $p(t, j)$. The default model has season-dependent transition probabilities and season- and survey-dependent detection probabilities, i.e. $\varphi(t)p(t, j)$, as used in MacKenzie et al. (2003). We show the `Maple` code with a small data set in Example

5.5 and then explore a real data set on house finches from MacKenzie et al. (2006).

Example 5.5 - Extrinsic parameter redundancy in multiple-seasons occupancy models: [See electronic appendix `example5.5.mw`] Consider an example where we have two seasons, with two surveys in the first season and three surveys in the second season. Suppose we only observe the occupancy-histories: 11|100, 10|100, 01|111, 00|011, 00|001 and 00|000. We can use the `Maple` procedure `occmodseasonshist` to obtain a model's parameter redundancy for this data set. The procedure `occmodseasonshist` requires two inputs to generate the model's exhaustive summary. The first input is a row vector of how many surveys there are per season, given in the vector $\mathbf{S}$ as

$$\mathbf{S} = \begin{bmatrix} 2 & 3 \end{bmatrix}.$$

The second input as a matrix of the observed data of the occupancy-histories, with the matrix being of dimension N as the number of occupancy-histories observed by $T \times \sum_{t=1}^{T} k_t$, given in the matrix $\mathbf{T}$ as

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consider a model which has season-dependent occupancy parameters and only survey-dependent detection parameters, as denoted by $\varphi(t)p(j)$. The parameter deficiency of the model can be found by the `Maple` code as follows:

```
S := < <2|3> >;
T := < <1|1|1|0|0>, <1|0|1|0|0>, <0|1|1|1|1>, <0|0|0|1|1>,
<0|0|0|0|1>, <0|0|0|0|0> >;
kappa := occmodseasonshist(2,3,S,T);
# Inputs: (y,z,S,T);
# y = transition probability; z = detection probability;
# for y and z: '1'=constant, '2'=season-dep, '3'=survey-dep,
#  '4'=season+survey-dep (y can only be '1' or '2');
# S = row vector of the number of surveys per season;
# T = list of all occupancy-histories.
```

```
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),pars):
r := Rank(DD); d := Dimension(theta)-r;
```

This `Maple` code generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_0 p_1 p_2 \varphi_{1,11} p_1 (1-p_2)(1-p_3) \\ \varphi_0 p_1 (1-p_2) \varphi_{1,11} p_1 (1-p_2)(1-p_3) \\ \varphi_0 (1-p_1) p_2 \varphi_{1,11} p_1 p_2 p_3 \\ \{\varphi_0 (1-p_1)(1-p_2)\varphi_{1,11} + (1-\varphi_0)\,\varphi_{1,1}\}\,(1-p_1) p_2 p_3 \\ \{\varphi_0 (1-p_1)(1-p_2)\varphi_{1,11} + (1-\varphi_0)\,\varphi_{1,1}\}\,(1-p_1)(1-p_2) p_3 \\ \varphi_0 (1-p_1)(1-p_2)\,\{\varphi_{1,11}(1-p_1)(1-p_2)(1-p_3) + (1-\varphi_{1,11})\} \cdots \\ \cdots + (1-\varphi_0)\,\{\varphi_{1,1}(1-p_1)(1-p_2)(1-p_3) + (1-\varphi_{1,1})\} \end{bmatrix},$$

where $\varphi_0 = \varphi_0^{[1]}$, $\varphi_{t,1} = \varphi_t^{[0,1]}$ and $\varphi_{t,11} = \varphi_t^{[1,1]}$ for season t. `Maple` does not work well with superscripts so we have used the subscript $t, 11$ to denote the superscript $[1, 1]$ for the occupancy parameter $\varphi_t^{[1,1]}$ where the species remains present between seasons, and $t, 1$ to denote the superscript $[0, 1]$ for the occupancy parameter $\varphi_t^{[0,1]}$ where the species goes from being absent to being present between seasons. The parameters in the model are $\boldsymbol{\theta} = [p_1,\ p_2,\ p_3,\ \varphi_0,\ \varphi_{1,1},\ \varphi_{1,11}]^T$, and differentiating the exhaustive summary by these parameters gives us the derivative matrix, which is not given here to save page space but can be found in the `Maple` file `example5.5.mw`. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{5}, \mathbf{d} := \mathbf{1}$.

This means that this exemplar data set has a parameter deficiency of 1, compared to the intrinsic parameter deficiency of 0 as proved in Theorem 5.1. □

We now use this `Maple` procedure to examine the extrinsic parameter redundancy of real data considering a data set of house finches from MacKenzie et al. (2006).

Example 5.6 - The house finches data set: This data set arises from a study of house finches (*Carpodacus mexicanus*) in North America. This study was conducted by the North American Breeding Bird Survey, which has been conducted annually since 1965 and uses observers who are volunteers. This data has been analysed before in Robbins et al. (1986) for the study between the years 1965 and 1979, in Sauer et al. (1994) which

details the effect of observer differences on inference, and in Link and Sauer (1998) which gives a location based analysis of the data. Further detailed information and historical background on the data set is given in Robbins et al. (1986) and MacKenzie et al. (2006). The volunteer observers stopped at 50 set points along a pre-determined route and listed all the different birds they could see in a three minute period of observation. We are only interested in house finches so the data set shows a '1' if a house finch was seen at one of the points and a '0' otherwise. In addition, a 'distance band' for each bird was recorded. These distance bands show us where each bird is seen from to their point of release from Long Island as the experimenters wished to explore the rate of house finches expanding westwards. These distance bands go from 0 to 2.5 kilometres in 0.1 kilometre increments and each observation been rounded to the nearest 0.1 kilometre increment. We use these distance band to explore certain sections of the data available to us, and there are more birds recorded for the lower distance bands than higher ones. These particular groupings below have no ecological interest and are only given for illustrative purposes.

- Data from all the distance bands (694 individual occupancy-histories).
- Data from the occupancy-histories with the distance band 0km (17 individual occupancy-histories).
- Data from the occupancy-histories with a distance band of between 2.2km and 2.5km (38 individual occupancy-histories).
- Data from the occupancy-histories with the distance band 2.5km (3 individual occupancy-histories).

Due to the data collection process, survey-dependent parameters are not possible in this analysis. Table 5.1 displays the possible occupancy models in this set-up, where the parameters could be constant $(\cdot)$ or season-dependent (t) for both the transition and detection probabilities. The number of parameters in the model are given in the second column and the parameter deficiency in the subsequent columns for each data subset specified above.

This is an example where our procedures have been used to find parameter deficiencies of certain collections of occupancy-histories rather than solely exploring intrinsic parameter deficiency for perfect data. The results in Table 5.1 show again the problem of insufficient data, as is a consistent theme in this thesis. What is somewhat surprising is that even though there are more occupancy-histories in the 2.2km to 2.5km group

Table 5.1: Extrinsic parameter redundancies in multiple-seasons occupancy models exploring the house finches data set from MacKenzie et al. (2006)

Model	Number of parameters	All bands deficiency	0km band deficiency	2.2-2.5km bands deficiency	2.5km band deficiency
$\varphi(\cdot)\ p(\cdot)$	4	0	0	0	2
$\varphi(\cdot)\ p(t)$	9	0	0	1	7
$\varphi(t)\ p(\cdot)$	12	0	1	3	10
$\varphi(t)\ p(t)$	17	0	1	5	15
Survey-dependence for detection probabilities is not possible here					

compared to the 0km distance band animals, the former case has a higher parameter redundancy. The reason for this is that there are quite a few histories in this group that have no detections at all (25 out of the 28) unlike the 0km group which has only 1 occupancy-history involving no detections at all during the study. This shows it is better for estimating parameters to have many distinct histories rather than solely many individual histories. □

5.7 Addition of group-dependent parameters

An extension to the multiple-seasons model could be to include a variety of groups of the species examined, as given in the **multiple-groups occupancy model** in this section. This is a model where different groups of the species can be differentiated between, such as species from different sites, or characteristics between species. In rich data sets it could also provide inference for different parameters between groups to analyse different group transition and detection probabilities.

One example of different group-dependent parameters is to differentiate between sampling sites. MacKenzie and Royle (2005), MacKenzie et al. (2006, Section 7.7), and Guillera-Arroita (2011) for example analyse the addition of more sites from the viewpoint of model design, and whether it is better to have more sites with fewer observations per site or to only have a few sites with a high number of observations. Other factors that could use this multiple-groups occupancy model would be to:

- Differentiate between breeding habitats between species.
- Differentiate between characteristics between the species, e.g. how big the animal is, if the animal has a particular biological feature or not, whether the species flowers or not, etc.

- Differentiate between a species' genetical information say by different ID tags.
- Differentiate between external influences such as the weather.

For example, we could consider the difference between male and female animals. If this was the case, the parameters could have the index $s = 1$ when the occupancy-history refers to male animals and $s = 2$ for female animals, so that the probability of detection of male animals is $p_{t,j,1}$ and the probability of detection of female animals is $p_{t,j,2}$ for survey j in season t. Note that when using this group-dependent model to account for these factors, the differences among species are then assumed to be *independent*. This can be quite unrealistic say in the case of gender where the number of male species is likely to be dependent on the number of female species due to such reasons as breeding. The different species case is a particular one we consider in Section 6.3 where we consider a two-species interaction model which takes into consideration non-independence between species.

The multiple-groups occupancy model has the following probability of occupancy-history h_X as

$$Pr(h_X) = \phi_{0,s}\left[\prod_{t=1}^{T-1}\Delta(\mathbf{p}_{X,t,s})\phi_{t,s}\right]\mathbf{p}_{X,T,s}, \tag{5.5}$$

with the same notation as before but considered at groups $s = 1, \ldots, S$. This means that the occupancy parameters are of the form $\varphi_{0,s}^{[1]}$, $\varphi_{t,s}^{[1,1]}$ and $\varphi_{t,s}^{[0,1]}$ with the detection parameters of the form $p_{t,j,s}$ for survey j in season t for group s. The overall likelihood is then equal to

$$L = \prod_{X=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study. In this case as occupancy-histories can observed from different groups, this could include multiple instances of the same occupancy-history observed at many groups. If there is there no group-dependency for all of the parameters, this model reduces to the multiple-seasons occupancy model of Equation (5.3) in Section 5.4. Additionally with this model, it is possible to consider group-dependence on only a certain set of parameters which gives this model added flexibility, e.g. the occupancy parameters may be group-dependent but the detection not or vice-versa depending on the study context.

In terms of intrinsic parameter redundancy for the multiple-groups occupancy model, the following Theorem 5.2 holds.

Theorem 5.2. *The multiple-groups occupancy model with perfect data with two or more surveys per season is* ***not*** *parameter redundant. Any season with only one survey or no surveys at all during the season for all groups adds the value of S to the parameter deficiency of that model.*

The proof of Theorem 5.2 is a direct result from Theorem 5.1 as if you consider intrinsic parameter redundancy separately for each group $s = 1, \ldots, S$, then the model has no parameter redundancy for each group s. As each individual group has no parameter redundancy, then all of the groups combined must also have no parameter redundancy as every parameter is group-dependent by the full model specification. The proof is also seen as a trivial application of a proof from integrated population modelling in Cole and McCrea (2012, Theorem 2: Remark 1).

The following `Maple` procedure `occmodgroupshist` can be used to examine group-dependent occupancy models, given in the supplementary `Maple` file `example5.7.mw`. This is an extension of the `Maple` procedure `occmodseasonshist` from Section 5.6. Due to group-dependence also being considered now, we have many more parameter options: The occupancy parameters can be constant, $\varphi(\cdot)$, season-dependent, $\varphi(t)$, and/or group-dependent, $\varphi(s)$. The detection parameters can be constant, $p(\cdot)$, survey-dependent, $p(j)$, season-dependent, $p(t)$, and/or group-dependent, $p(s)$, i.e. in total there are 32 different possibilities of model. We show a small exemplar data set of occupancy-histories to show how the procedure works, then explore a real data set of amphibian monitoring in the Greater Yellowstone and Grand Teton national parks.

Example 5.7 - Extrinsic parameter redundancy in the multiple-groups occupancy model: [See electronic appendix `example5.7.mw`] Consider an example with two seasons where there are two surveys in both seasons. There are two different groups in this example and we observe only the occupancy-histories 11|10(1), 11|01(1), 01|10(1), 01|00(1), 00|10(1), 00|00(1), 11|01(2), 01|00(2), 00|10(2), with the number in brackets indicating at what group the occupancy-history was observed at.

For this procedure, we require three data inputs. The first input in the row vector **S** is the number of surveys per season, given as

$$\mathbf{S} = \begin{bmatrix} 2 & 2 \end{bmatrix}.$$

The second input then specifies which occupancy-history is in each group as the column

vector **S2** as

$$\mathbf{S2} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \end{bmatrix}^T.$$

The final input is the matrix $\mathbf{T}$ of all the occupancy-histories as

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that each occupancy-history in row i in the matrix $\mathbf{T}$ was observed at the group in row i of the column vector **S2** respectively. Consider a model where the occupancy parameters are only season-dependent and the detection parameters are only survey- and group-dependent. The parameter deficiency of these occupancy-histories can then be found by the following `Maple` code over.

```
S := < <2|2> >;
S2 := < <1>, <1>, <1>, <1>, <1>, <1>, <2>, <2>, <2>, <2> >;
T := < <1|1|1|0>, <1|1|0|1>, <0|1|1|0>, <0|1|0|0>, <0|0|1|0>,
       <0|0|0|0>, <1|1|0|1>, <0|1|0|0>, <0|0|1|0>, <0|0|0|0> >;
kappa := occmodgroupshist(2,7,S,S2,T);
# Inputs: (y,z,S,S2,T);
# y = transition probability; z = detection probability;
# for y: ‘1’=constant, ‘2’=season-dep, ‘3’=group-dep,
# ‘4’=season+group-dep;
# for z: ‘1’=constant, ‘2’=season-dep, ‘3’=survey-dep,
# ‘4’=season+survey-dep; ‘5’=group-dep, ‘6’=season+group-dep,
# ‘7’=survey+group-dep, ‘8’=season+survey+group-dep;
# S = row vector of the number of surveys per season;
# S2 = column vector determining the group of the occupancy-history;
# T = list of all occupancy-histories.
```

```
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),theta):
r := Rank(DD); d := Dimension(theta)-r;
```

This generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_{0,c}p_{c,1,1}p_{c,2,1}\varphi_{1,11,c}p_{c,1,1}(1-p_{c,2,1}) \\ \varphi_{0,c}p_{c,1,1}p_{c,2,1}\varphi_{1,11,c}(1-p_{c,1,1})p_{c,2,1} \\ \varphi_{0,c}(1-p_{c,1,1})p_{c,2,1}\varphi_{1,11,c}p_{c,1,1}(1-p_{c,2,1}) \\ \varphi_{0,c}(1-p_{c,1,1})p_{c,2,1}\left\{\varphi_{1,11,c}(1-p_{c,1,1})(1-p_{c,2,1})+(1-\varphi_{1,11,c})\right\} \\ \varphi_{0,c}(1-p_{c,1,1})(1-p_{c,2,1})\varphi_{1,11,c}p_{c,1,1}(1-p_{c,2,1})+(1-\varphi_{0,c})\varphi_{1,1,c}p_{c,1,1}(1-p_{c,2,1}) \\ \varphi_{0,c}(1-p_{c,1,1})(1-p_{c,2,1})\left\{\varphi_{1,11,c}(1-p_{c,1,1})(1-p_{c,2,1})+(1-\varphi_{1,11,c})\right\}+\cdots \\ \cdots(1-\varphi_{0,c})\left\{\varphi_{1,1,c}(1-p_{c,1,1})(1-p_{c,2,1})+(1-\varphi_{1,1,c})\right\} \\ \varphi_{0,c}p_{c,1,2}p_{c,2,2}\varphi_{1,11,c}(1-p_{c,1,2})p_{c,2,2} \\ \varphi_{0,c}(1-p_{c,1,2})p_{c,2,2}\left\{\varphi_{1,11,c}(1-p_{c,1,2})(1-p_{c,2,2})+(1-\varphi_{1,11,c})\right\} \\ \varphi_{0,c}(1-p_{c,1,2})(1-p_{c,2,2})\varphi_{1,11,c}p_{c,1,2}(1-p_{c,2,2})+(1-\varphi_{0,c})\varphi_{1,1,c}p_{c,1,2}(1-p_{c,2,2}) \\ \varphi_{0,c}(1-p_{c,1,2})(1-p_{c,2,2})\left\{\varphi_{1,11,c}(1-p_{c,1,2})(1-p_{c,2,2})+(1-\varphi_{1,11,c})\right\}+\cdots \\ \cdots(1-\varphi_{0,c})\left\{\varphi_{1,1,c}(1-p_{c,1,2})(1-p_{c,2,2})+(1-\varphi_{1,1,c})\right\} \end{bmatrix},$$

where $\varphi_{0,c} = \varphi_{0,c}^{[1]}$, $\varphi_{1,1,c} = \varphi_{1,c}^{[0,1]}$ and $\varphi_{1,11,c} = \varphi_{1,c}^{[1,1]}$ indicates how they are represented in `Maple`. The parameters in the model are $\boldsymbol{\theta} = [p_{c,1,1}, p_{c,1,2}, p_{c,2,1}, p_{c,2,2}, \varphi_{0,c}, \varphi_{1,1,c}, \varphi_{1,11,c}]^T$, and differentiating the exhaustive summary by these parameters gives us the derivative matrix, which is not given here to save page space but can be found in the `Maple` file `example5.7.mw`. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{7}, \mathbf{d} := \mathbf{0}$.

In this model there are sufficient data to theoretically allow all the parameters in the model to be estimated with no parameter redundancy. □

We now examine extrinsic parameter redundancy in a real data set involving monitoring amphibians in the Yellowstone and Grand Teton National Parks in the USA.

Example 5.8 - The Yellowstone and Grand Teton amphibians data set: This data set arises from the monitoring of amphibians in the Yellowstone and Grand Teton National Parks in the state of Wyoming, USA, over a four-year period from 2006 to 2009.

Gould et al. (2012) derives occupancy estimates of amphibians from data gathered at two different 'scales', one from surveys of the portions of watersheds denoted as catchments, and another scale from surveys of individual wetland sites. There are also three different species of amphibian considered, the tiger salamander (*Ambystoma tigrinum*), the Columbia spotted frog (*Rana luteiventris*), and the boreal chorus frog (*Pseudacris maculata*). We consider a multiple species interaction model in Section 6.3 but we assume there is independence between the three species in the data set for this analysis. The scale of the surveys which are taken from individual wetland sites is either recorded at Yellowstone or Grand Teton National Park, and the surveys which are not taken at wetland sites are from one of 31 different catchment sites in the study. The parameters could be season-dependent (t) and/or group-dependent (s), and the detection parameters additionally could be survey-dependent (j).

We explore group-dependent models by examining arbitrary subsets of the data to illustrate typical parameter deficiencies for a range of sparse and rich data sets. The three different group-dependent subsets we consider in Table 5.2 below are

- Data dependent on surveys from the individual wetland sites only (a total of 78 occupancy-histories with either Yellowstone or Grand Teton as the different groups).
- Data dependent on surveys from catchments and only considering the first 9 sites (a total of 504 occupancy-histories with 9 different groups).
- Data dependent on surveys from catchments and only considering the first 18 sites (a total of 663 occupancy-histories with 18 different groups).

The subsets given above are in terms of the data's richness, with the wetland-only data being relatively rich as there are many distinct occupancy-histories with only two different groups, while the subset considering 18 different groups is sparse due to the large number of parameters when they are considered to be group-dependent. We further eliminate all of the occupancy-histories which have missing values; We deal with missing observations in Section 5.8. The results of this extrinsic parameter redundancy analysis are displayed in Table 5.2 which gives the model in the first column and each model rank and parameter deficiency for the three partitions above in the remaining columns.

As we can see from Table 5.2 that there are extremely high parameter deficiencies for the most complex occupancy models where there is quite sparse data. The reason

Table 5.2: Extrinsic parameter redundancies in multiple-groups occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012) where all histories with missing observations are ignored

Factor:	Wetland sites		Site ID for 9 sites		Site ID for 18 sites	
Model	Rank	Deficiency	Rank	Deficiency	Rank	Deficiency
$\varphi(\cdot)\ p(\cdot)$	4	0	4	0	4	0
$\varphi(\cdot)\ p(t)$	7	0	7	0	7	0
$\varphi(\cdot)\ p(j)$	5	0	5	0	5	0
$\varphi(\cdot)\ p(t,j)$	11	0	11	0	11	0
$\varphi(\cdot)\ p(s)$	5	0	12	0	21	0
$\varphi(\cdot)\ p(t,s)$	11	0	39	0	66	9
$\varphi(\cdot)\ p(j,s)$	7	0	21	0	38	1
$\varphi(\cdot)\ p(t,j,s)$	19	0	74	1	106	41
$\varphi(t)\ p(\cdot)$	8	0	8	0	8	0
$\varphi(t)\ p(t)$	11	0	11	0	11	0
$\varphi(t)\ p(j)$	9	0	9	0	9	0
$\varphi(t)\ p(t,j)$	15	0	15	0	15	0
$\varphi(t)\ p(s)$	9	0	16	0	25	0
$\varphi(t)\ p(t,s)$	15	0	43	0	70	9
$\varphi(t)\ p(j,s)$	11	0	25	0	42	1
$\varphi(t)\ p(t,j,s)$	23	0	78	1	110	41
$\varphi(s)\ p(\cdot)$	7	0	28	0	50	5
$\varphi(s)\ p(t)$	10	0	31	0	53	5
$\varphi(s)\ p(j)$	8	0	29	0	51	5
$\varphi(s)\ p(t,j)$	14	0	35	0	57	5
$\varphi(s)\ p(s)$	8	0	36	0	63	9
$\varphi(s)\ p(t,s)$	14	0	63	0	95	31
$\varphi(s)\ p(j,s)$	10	0	45	0	75	15
$\varphi(s)\ p(t,j,s)$	20	2	92	7	124	74
$\varphi(t,s)\ p(\cdot)$	15	0	62	2	91	36
$\varphi(t,s)\ p(t)$	18	0	65	2	94	36
$\varphi(t,s)\ p(j)$	16	0	63	2	92	36
$\varphi(t,s)\ p(t,j)$	22	0	69	2	98	36
$\varphi(t,s)\ p(s)$	16	0	70	2	101	43
$\varphi(t,s)\ p(t,s)$	20	2	91	8	123	75
$\varphi(t,s)\ p(j,s)$	17	1	78	3	110	52
$\varphi(t,s)\ p(t,j,s)$	24	6	107	28	139	131

for this is the high number of different parameters in the complex models and the lack of data: The data needs to be very rich for there to be no parameter redundancy in the $\varphi(t,s)p(t,j,s)$ model for example. A number of the simpler models turn out to have no parameter redundancy though, such as all of the models which do not include group-dependence in any of the parameters in the model. □

5.8 Missing observations in occupancy models

We have explored the Yellowstone and Grand Teton national park data set of Gould et al. (2012) in the previous section, but a limitation of our analysis is that we had to delete some of the data as there were missing observations in some of the occupancy-histories. In the Yellowstone and Grand Teton data set we actually discarded 519 of the 1290 possible occupancy-histories as there were some missing observations in them, even though these 519 occupancy-histories still provide some information about the species. This is a poor way of dealing with such incomplete histories as we should maximise the data given to us fully, and not discard these partial occupancy-histories. Leaving out data may also bias our parameter estimates we obtain from the data as we are discarding some occupancy-histories that contain additional information about some of the model parameters. Missing observations have been dealt with in occupancy models in MacKenzie et al. (2002, 2003, 2004, 2009), and it is relatively easy to adjust the model accordingly for these missing observations. We regard this as a very practical consideration as quite a number of the data sets we have analysed using occupancy models have missing observations in them. The reasons for these missing observations are wide-ranging, from a lack of resources to technological errors, though it is important that we are considering missing observations rather than *incorrect* observations. We show how missing observations can be considered in a general model first, though it is easier to follow the adjustments for missing observations by using examples as seen in Examples 5.9 to 5.12.

We maintain the previous notation of $\varphi_{0,s}^{[1]}$ as the original occupancy state in the study, $\varphi_{t,s}^{[0,1]}$ and $\varphi_{t,s}^{[1,1]}$ as the occupancy transition probabilities, along with $p_{t,j,s}$ as the detection probabilities. All these parameters are for seasons $t = 1, \ldots, T$ at groups $s = 1, \ldots, S$, and the detection probabilities are for every survey $j = 1, \ldots, k_t$ where k_t denotes the number of surveys in season t for every group s. The transition occupancy matrices remain as

$$\phi_{t,s} = \begin{bmatrix} \varphi_{t,s}^{[1,1]} & 1 - \varphi_{t,s}^{[1,1]} \\ \varphi_{t,s}^{[0,1]} & 1 - \varphi_{t,s}^{[0,1]} \end{bmatrix},$$

for all $t = 1, \ldots, T-1$ at groups $s = 1, \ldots, S$, where $\phi_{0,s}$ is equal to the row vector

$$\phi_{0,s} = \left[\begin{array}{cc} \varphi_{0,s}^{[1]} & 1 - \varphi_{0,s}^{[1]} \end{array} \right].$$

Let the observed state in an occupancy-history in survey j in season t at group s be denoted by $\delta_{t,j,s}$. If we further let $\gamma_{t,j,s}$ be defined as

$$\gamma_{t,j,s} = \begin{cases} 1 & \text{when } \delta_{t,j,s} \geq 0 \text{ (i.e. it is not a missing observation),} \\ 0 & \text{when } \delta_{t,j,s} = -1 \text{ (i.e. it } \textit{is} \text{ a missing observation),} \end{cases}$$

then $\mathbf{p}_{X,t,s}$ is the column vector

$$\mathbf{p}_{X,t,s} = \left[\begin{array}{c} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \delta_{t,j,s} p_{t,j,s} + (1 - \delta_{t,j,s}) (1 - p_{t,j,s}) \right\} + (1 - \gamma_{t,j,s}) \\ \prod_{j=1}^{k_t} (1 - \delta_{t,j,s} \gamma_{t,j,s}) \end{array} \right]. \tag{5.6}$$

Observe the differences between the $\mathbf{p}$ vectors in Equations (5.2) and (5.6) and how an adjustment is made for missing observations using the $\gamma_{t,j,s}$ terms. The second term in $\mathbf{p}_{X,t,s}$ can also be represented as

$$\begin{cases} 0 & \text{when a single } \delta_{t,j,s} = 1 \text{ for any survey } j \text{ in that season for that group,} \\ 1 & \text{when } \delta_{t,j,s} = \{0, -1\} \text{ for all surveys in that season for that group.} \end{cases}$$

We can again denote $\Delta(\mathbf{p}_{X,t})$ to be a diagonal matrix with diagonal entries given by the vector $\mathbf{p}_{X,t}$, and zeros on the non-diagonal entries. The probability of an occupancy-history, h_X, then remains

$$Pr(h_X) = \phi_{0,s} \left[\prod_{t=1}^{T-1} \Delta(\mathbf{p}_{X,t,s}) \phi_{t,s} \right] \mathbf{p}_{X,T,s},$$

with the final likelihood

$$L = \prod_{X=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study.

This model which adapts for missing observations has been implemented in the previous `Maple` procedures of `occmodseasonshist` from Section 5.6 and `occmodgroupshist` of Section 5.7. We denote an occupancy-history which has a missing observation with an underline, '_', such as in the occupancy-history $h = 11|1\_|00$ in Example 5.9 below, where the detection for the second survey of the second season was missing. When inputting occupancy-histories into `Maple` we use the number '-1' to denote when that particular survey has a missing observation. We now demonstrate in Examples 5.9 to 5.12 how some occupancy-histories can be displayed as their probabilities of being observed.

Example 5.9: Suppose we observe the occupancy-history $h = 11|1\_|00$ where the detection for the second survey of the second season was missing. It would then be incorrect to have any $p_{2,2}$ parameters in this occupancy-history as the occupancy-history gives us no information about the probability of detection for the second survey of the second season. If we assume that there is no for that group-dependence for ease of notation here, then the probability of this occupancy-history is equal to

$$\begin{aligned} Pr(h) &= \phi_0 \Delta(\mathbf{p}_{11,1})\phi_1 \Delta(\mathbf{p}_{1\_,2})\phi_2 \mathbf{p}_{00,3} \\ &= \begin{bmatrix} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}p_{1,2} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} p_{2,1} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1})(1-p_{3,2}) \\ 1 \end{bmatrix} \\ &= \varphi_0^{[1]} p_{1,1} p_{1,2} \varphi_1^{[1,1]} p_{2,1} \left\{ \varphi_2^{[1,1]} (1-p_{3,1})(1-p_{3,2}) + (1-\varphi_2^{[1,1]}) \right\}. \end{aligned}$$

We can see here that the second survey in the second season is effectively overlooked in the occupancy-history so that there are no $p_{2,2}$ parameters in the model. □

Example 5.10: Suppose we observe the occupancy-history $h = 11|0\_|00$ which is the same occupancy-history at Example 5.9 apart from there is no detection in the first survey of the second season. If there is no for that group-dependence, the probability

of this occupancy-history is given as

$$\begin{aligned}
Pr(h) &= \phi_0\Delta(\mathbf{p}_{11,1})\phi_1\Delta(\mathbf{p}_{0_,2})\phi_2\mathbf{p}_{00,3} \\
&= \begin{bmatrix} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}p_{1,2} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} (1-p_{2,1}) & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1})(1-p_{3,2}) \\ 1 \end{bmatrix} \\
&= \varphi_0^{[1]}p_{1,1}p_{1,2}\left[\varphi_1^{[1,1]}(1-p_{2,1})\left\{\varphi_2^{[1,1]}(1-p_{3,1})(1-p_{3,2})+(1-\varphi_2^{[1,1]})\right\}+\right. \\
&\quad \left.(1-\varphi_1^{[1,1]})\left\{\varphi_2^{[0,1]}(1-p_{3,1})(1-p_{3,2})+(1-\varphi_2^{[0,1]})\right\}\right].
\end{aligned}$$

There is uncertainty of what occupancy state the species was in during the second season as it still could be absent or present but they were not detected. □

Example 5.11: A further problem is if there are no detections at all in a single season. Take the occupancy-history $h = 11|\_\_|00$ which has this problem where no detection surveys were possible in the second season. Then similarly to above, the season is effectively ignored for the purposes of the occupancy-history. However some adjustment does need to be made to consider the occupancy state in this unknown season as we are unsure if the species was present or absent during the season. The probability of this occupancy-history then is equal to

$$\begin{aligned}
Pr(h) &= \phi_0\Delta(\mathbf{p}_{11,1})\phi_1\Delta(\mathbf{p}_{__,2})\phi_2\mathbf{p}_{00,3} \\
&= \begin{bmatrix} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}p_{1,2} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1})(1-p_{3,2}) \\ 1 \end{bmatrix} \\
&= \varphi_0^{[1]}p_{1,1}p_{1,2}\varphi_1^{[1,1]}\left\{\varphi_2^{[1,1]}(1-p_{3,1})(1-p_{3,2})+(1-\varphi_2^{[1,1]})\right\}+ \\
&\quad \varphi_0^{[1]}p_{1,1}p_{1,2}\left(1-\varphi_1^{[1,1]}\right)\left\{\varphi_2^{[0,1]}(1-p_{3,1})(1-p_{3,2})+(1-\varphi_2^{[0,1]})\right\}.
\end{aligned}$$

Therefore in the case where there are no detections in a single season, $\mathbf{p}_{\_\_,t} = [1\ 1]^T$ and $\Delta(\mathbf{p}_{\_\_,t})$ is equal to the identity matrix. □

Example 5.12: Further note that if all of the detections in the final season of the study are missing, then the parameters $\varphi_{T-1}^{[1,1]}$ and $\varphi_{T-1}^{[0,1]}$ are not given in the exhaustive

summary term. This can be seen in the example $h = 11|01|\_\_$:

$$\begin{aligned} Pr(h) &= \phi_0\Delta(\mathbf{p}_{11,1})\phi_1\Delta(\mathbf{p}_{01,2})\phi_2\mathbf{p}_{\_\_,3} \\ &= \left[\begin{array}{cc} \varphi_0^{[1]} & 1-\varphi_0^{[1]} \end{array}\right] \cdot \left[\begin{array}{cc} p_{1,1}p_{1,2} & 0 \\ 0 & 0 \end{array}\right] \cdot \left[\begin{array}{cc} \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]} \\ \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]} \end{array}\right] \cdot \\ &\quad \left[\begin{array}{cc} (1-p_{2,1})p_{2,2} & 0 \\ 0 & 0 \end{array}\right] \cdot \left[\begin{array}{cc} \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]} \\ \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]} \end{array}\right] \cdot \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \\ &= \varphi_0^{[1]}p_{1,1}p_{1,2}\varphi_1^{[1,1]}(1-p_{2,1})p_{2,2}\left\{\varphi_2^{[1,1]} + (1-\varphi_2^{[1,1]})\right\} \\ &= \varphi_0^{[1]}p_{1,1}p_{1,2}\varphi_1^{[1,1]}(1-p_{2,1})p_{2,2}. \end{aligned}$$

This idea that the parameters $\varphi_{T-1}^{[1,1]}$ and $\varphi_{T-1}^{[0,1]}$ cannot be estimated further extends to the case where more than one season has no surveys: If for example the occupancy-history $11|\_\_|\_\_$ is given, if the result is simplified, then the probability of this occupancy-history is equal to $\varphi_0^{[1]}p_{1,1}p_{1,2}$ with no $\varphi_t^{[a,1]}$ parameters estimable. □

Example 5.8 revisited - Missing observations considered in the Yellowstone and Grand Teton data set: We show how parameter redundancy can be reduced from Table 5.2 by now considering the occupancy-histories with missing observations in the analysis. We give a reminder of how the different subsets of the data have been differentiated below:

- Data dependent on surveys from the individual wetland sites only (a total of 78 occupancy-histories with either Yellowstone or Grand Teton as the different groups).
- Data dependent on surveys from catchments and only considering the first 9 sites (a total of 504 occupancy-histories with 9 different groups).
- Data dependent on surveys from catchments and only considering the first 18 sites (a total of 663 occupancy-histories with 18 different groups).

The models are denoted the same as given in Table 5.2 previously in this chapter. Table 5.3 shows these extrinsic parameter redundancy results when these incomplete histories are then considered in the analysis.

We can see by comparing Table 5.2 to Table 5.3 that there is less parameter deficiency when the model is parameter redundant. There are also some models which then become full rank only after these missing observation occupancy-histories are

Table 5.3: Extrinsic parameter redundancies in multiple-groups occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012) where all histories with missing observations are now included

Factor:	Wetland sites		Site ID for 9 sites		Site ID for 18 sites	
Model	Rank	Deficiency	Rank	Deficiency	Rank	Deficiency
$\varphi(\cdot)\ p(\cdot)$	4	0	4	0	4	0
$\varphi(\cdot)\ p(t)$	7	0	7	0	7	0
$\varphi(\cdot)\ p(j)$	5	0	5	0	5	0
$\varphi(\cdot)\ p(t,j)$	11	0	11	0	11	0
$\varphi(\cdot)\ p(s)$	5	0	12	0	21	0
$\varphi(\cdot)\ p(t,s)$	11	0	39	0	*74*	*1*
$\varphi(\cdot)\ p(j,s)$	7	0	21	0	38	1
$\varphi(\cdot)\ p(t,j,s)$	19	0	**75**	**0**	*141*	*6*
$\varphi(t)\ p(\cdot)$	8	0	8	0	8	0
$\varphi(t)\ p(t)$	11	0	11	0	11	0
$\varphi(t)\ p(j)$	9	0	9	0	9	0
$\varphi(t)\ p(t,j)$	15	0	15	0	15	0
$\varphi(t)\ p(s)$	9	0	16	0	25	0
$\varphi(t)\ p(t,s)$	15	0	43	0	*78*	*1*
$\varphi(t)\ p(j,s)$	11	0	25	0	42	1
$\varphi(t)\ p(t,j,s)$	23	0	**79**	**0**	*145*	*6*
$\varphi(s)\ p(\cdot)$	7	0	28	0	**55**	**0**
$\varphi(s)\ p(t)$	10	0	31	0	**58**	**0**
$\varphi(s)\ p(j)$	8	0	29	0	**56**	**0**
$\varphi(s)\ p(t,j)$	14	0	35	0	**62**	**0**
$\varphi(s)\ p(s)$	8	0	36	0	*71*	*1*
$\varphi(s)\ p(t,s)$	14	0	63	0	*122*	*4*
$\varphi(s)\ p(j,s)$	10	0	45	0	*88*	*2*
$\varphi(s)\ p(t,j,s)$	**22**	**0**	**99**	**0**	*183*	*15*
$\varphi(t,s)\ p(\cdot)$	15	0	**64**	**0**	*123*	*4*
$\varphi(t,s)\ p(t)$	18	0	**67**	**0**	*126*	*4*
$\varphi(t,s)\ p(j)$	16	0	**65**	**0**	*124*	*4*
$\varphi(t,s)\ p(t,j)$	22	0	**71**	**0**	*130*	*4*
$\varphi(t,s)\ p(s)$	16	0	**72**	**0**	*139*	*5*
$\varphi(t,s)\ p(t,s)$	**22**	**0**	**99**	**0**	*183*	*15*
$\varphi(t,s)\ p(j,s)$	**18**	**0**	**81**	**0**	*155*	*7*
$\varphi(t,s)\ p(t,j,s)$	*26*	*4*	*134*	*1*	*230*	*40*

The **bold** values indicate that a model which was parameter redundant when missing observations were not considered in Table 5.2 is now not parameter redundant.

The *italics* values indicate there is a reduction in the parameter deficiency for that model compared to the data excluding missing observations in Table 5.2.

considered, such as the $\varphi(t)p(j,s)$, $\varphi(s)p(t)$ and $\varphi(s)p(t,j)$ occupancy models for the 18-site data set. This shows that it is beneficial to use incomplete missing data as more parameters can be estimated if we consider these occupancy-histories. It is also realistic in studies that some of the observations will be missing so that it is important that these occupancy-histories are considered in the analysis. □

5.9 The effect of data sparseness on parameter redundancy

As in Sections 3.8 and 4.7, we would like to identify how parameter redundancy increases with sparseness of data. We explore combinations of low, medium and high occupancy and detection parameters and see how parameter redundant they are when we have sparse data. We view cases where there are only $N = 20$, 30, 50 or 100 occupancy-histories observed in the study. As you can see in Table 5.4 on page 174, the probabilities of recording an occupancy-history where there are detections in every survey is especially small, even for the case where occupancy and detection probabilities are both relatively high (0.8 is defined as a 'high' value in this example).

We consider for given parameter values and sample sizes the expected probability of each occupancy-history. If the expectation is at least 1, then the occupancy-history is included in the exhaustive summary, otherwise the occupancy-history is excluded from the exhaustive summary. We then obtain the parameter redundancy of the exhaustive summary where only the occupancy-histories with expectation greater than 1 are included. For example, consider a 2-season model with 3-surveys per season when the occupancy and detection parameters are equal to $\varphi = p = 0.2$ (which is defined as 'low' in this example). When $N = 20$ there is only one occupancy-history with an expectation greater than one in this case, which is the history 000|000. However, when $\varphi = 0.2$ and $p = 0.8$, there are three different occupancy-histories with expectations greater than one as 000|000, 000|111 and 111|000. Furthermore, if $\varphi = 0.8$ and $p = 0.2$, there are then seven different occupancy-histories with expectations greater than one as 000|000, 100|000, 010|000, 001|000, 000|100, 000|010, 000|001. We can increase N to see how this increases the number of occupancy-histories which have greater expectations than one and how this affects the parameter redundancy in the model. We can also list how large N is required to be so that each occupancy model is full rank. The results of this are shown in Table 5.5 on page 175 for a 2-season 3-survey model and in Tables 5.6a and 5.6b on pages 176 and 177 for a 3-season 2-survey model (note

that season-dependent occupancy parameters are only applicable when there are three seasons in the study).

We can see in Tables 5.5, 5.6a and 5.6b that there is generally less parameter redundancy if the detection probabilities are quite high as there are more occupancy-histories with expectations greater than one when this occurs. A similar effect occurs when occupancy probabilities are also quite high. The real problems with parameter redundancy are for data sets where detection and occupation parameters are both low; many occupancy-histories need to be recorded in this case so that all the parameters in the model can be estimated.

Table 5.4: The probabilities of occupancy-histories being observed for a certain number of detections during the history, with different ranges of low, medium and high occupancy and detection probabilities

Number of	$\varphi = 0.2$			$\varphi = 0.5$			$\varphi = 0.8$		
detections	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$
The 2-seasons with 3-surveys per season model (Table 5.5):									
0	0.8143	0.6806	0.6426	0.5715	0.3164	0.2540	0.3716	0.0900	0.0426
1	0.1386	0.1238	0.0308	0.2903	0.2109	0.0484	0.3745	0.1800	0.0317
2	0.0406	0.1294	0.1235	0.1094	0.2461	0.1958	0.1880	0.2700	0.1327
3	0.0058	0.0525	0.1671	0.0245	0.1406	0.2765	0.0550	0.2400	0.2163
4	0.0006	0.0094	0.0098	0.0038	0.0586	0.0614	0.0098	0.1500	0.1573
5	0.0001	0.0038	0.0157	0.0004	0.0234	0.0983	0.0010	0.0600	0.2517
6	0.0000	0.0006	0.0105	0.0000	0.0039	0.0655	0.0000	0.0100	0.1678
The 3-seasons with 2-surveys per season model (Tables 5.6a and 5.6b):									
0	0.7992	0.6141	0.5275	0.5514	0.2441	0.1406	0.3609	0.0640	0.0125
1	0.1654	0.2168	0.1254	0.3228	0.2930	0.1298	0.3893	0.1920	0.0413
2	0.0321	0.1339	0.2606	0.1033	0.2637	0.2995	0.1887	0.2880	0.1283
3	0.0031	0.0265	0.0400	0.0198	0.1328	0.1638	0.0518	0.2560	0.1992
4	0.0003	0.0079	0.0413	0.0025	0.0527	0.1843	0.0085	0.1440	0.2831
5	0.0000	0.0008	0.0032	0.0002	0.0117	0.0492	0.0008	0.0480	0.2013
6	0.0000	0.0001	0.0021	0.0000	0.0020	0.0328	0.0000	0.0080	0.1342
The first column gives the number of detections in a single occupancy-history. The subsequent columns give the probabilities of that number of detections being observed for the different values of φ and p given in the top row.									

Table 5.5: Parameter redundancies for the occupancy-histories in a 2-season 3-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities

	$\varphi = 0.2$			$\varphi = 0.5$			$\varphi = 0.8$		
	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$
$N = 20$									
# of histories	1	1	3	1	1	4	7	1	3
$\varphi\ p(\cdot)$	3	3	1	3	3	0	1	3	1
$\varphi\ p(t)$	4	4	2	4	4	1	2	4	2
$\varphi\ p(j)$	5	5	3	5	5	2	1	5	3
$\varphi\ p(t,j)$	8	8	6	8	8	5	2	8	6
$N = 30$									
# of histories	1	1	3	7	15	4	7	1	10
$\varphi\ p(\cdot)$	3	3	1	1	0	0	1	3	0
$\varphi\ p(t)$	4	4	2	2	0	1	2	4	0
$\varphi\ p(j)$	5	5	3	1	0	2	1	5	0
$\varphi\ p(t,j)$	8	8	6	2	0	5	2	8	0
$N = 50$									
# of histories	7	15	9	7	15	10	7	15	16
$\varphi\ p(\cdot)$	1	0	0	1	0	0	1	0	0
$\varphi\ p(t)$	2	0	0	2	0	0	2	0	0
$\varphi\ p(j)$	1	0	0	1	0	0	1	0	0
$\varphi\ p(t,j)$	2	0	0	2	0	0	2	0	0
$N = 100$									
# of histories	7	15	10	13	15	16	22	64	31
$\varphi\ p(\cdot)$	1	0	0	0	0	0	0	0	0
$\varphi\ p(t)$	2	0	0	0	0	0	0	0	0
$\varphi\ p(j)$	1	0	0	0	0	0	0	0	0
$\varphi\ p(t,j)$	2	0	0	0	0	0	0	0	0
How many observations are needed to ensure every model is full rank:									
	174	49	49	83	29	32	65	34	24
# of histories indicates how many occupancy-histories have an expected value greater than one.									

Table 5.6a: Parameter redundancies for the occupancy-histories in a 3-season 2-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities (Table 1)

	$\varphi = 0.2$			$\varphi = 0.5$			$\varphi = 0.8$		
	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$
$N = 20$									
# of histories	1	1	4	7	1	7	7	1	4
$\varphi(\cdot)\ p(\cdot)$	3	3	0	0	3	0	0	3	0
$\varphi(\cdot)\ p(t)$	5	5	2	2	5	0	2	5	2
$\varphi(\cdot)\ p(j)$	4	4	1	0	4	0	0	4	1
$\varphi(\cdot)\ p(t,j)$	8	8	5	2	8	2	2	8	5
$\varphi(t)\ p(\cdot)$	5	5	2	2	5	0	2	5	2
$\varphi(t)\ p(t)$	7	7	4	4	7	1	4	7	4
$\varphi(t)\ p(j)$	6	6	3	2	6	0	2	6	3
$\varphi(t)\ p(t,j)$	10	10	7	4	10	4	4	10	7
$N = 30$									
# of histories	1	10	4	7	10	7	7	1	10
$\varphi(\cdot)\ p(\cdot)$	3	0	0	0	0	0	0	3	0
$\varphi(\cdot)\ p(t)$	5	0	2	2	0	0	2	5	0
$\varphi(\cdot)\ p(j)$	4	0	1	0	0	0	0	4	0
$\varphi(\cdot)\ p(t,j)$	8	0	5	2	0	2	2	8	0
$\varphi(t)\ p(\cdot)$	5	1	2	2	1	0	2	5	1
$\varphi(t)\ p(t)$	7	1	4	4	1	1	4	7	1
$\varphi(t)\ p(j)$	6	1	3	2	1	0	2	6	1
$\varphi(t)\ p(t,j)$	10	1	7	4	1	4	4	10	1
# of histories indicates how many occupancy-histories have an expected value greater than one.									

Table 5.6b: Parameter redundancies for the occupancy-histories in a 3-season 2-surveys occupancy model with different ranges of low/medium/high occupancy and detection probabilities (Table 2)

	$\varphi = 0.2$			$\varphi = 0.5$			$\varphi = 0.8$		
	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$	$p = 0.2$	$p = 0.5$	$p = 0.8$
$N = 50$									
# of histories	7	10	10	7	10	14	7	10	13
$\varphi(\cdot)\ p(\cdot)$	0	0	0	0	0	0	0	0	0
$\varphi(\cdot)\ p(t)$	2	0	0	2	0	0	2	0	0
$\varphi(\cdot)\ p(j)$	0	0	0	0	0	0	0	0	0
$\varphi(\cdot)\ p(t,j)$	2	0	0	2	0	0	2	0	0
$\varphi(t)\ p(\cdot)$	2	1	1	2	1	0	2	1	0
$\varphi(t)\ p(t)$	4	1	1	4	1	0	4	1	0
$\varphi(t)\ p(j)$	2	1	1	2	1	0	2	1	0
$\varphi(t)\ p(t,j)$	4	1	1	4	1	0	4	1	0
$N = 100$									
# of histories	7	10	13	10	10	26	22	37	26
$\varphi(\cdot)\ p(\cdot)$	0	0	0	0	0	0	0	0	0
$\varphi(\cdot)\ p(t)$	2	0	0	0	0	0	0	0	0
$\varphi(\cdot)\ p(j)$	0	0	0	0	0	0	0	0	0
$\varphi(\cdot)\ p(t,j)$	2	0	0	0	0	0	0	0	0
$\varphi(t)\ p(\cdot)$	2	1	0	1	1	0	0	0	0
$\varphi(t)\ p(t)$	4	1	0	1	1	0	0	1	0
$\varphi(t)\ p(j)$	2	1	0	1	1	0	0	0	0
$\varphi(t)\ p(t,j)$	4	1	0	1	1	0	0	0	0
How many observations are needed to ensure every $\varphi(\cdot)$ model is full rank:									
	146	28	48	75	21	47	62	32	30
How many observations are needed to ensure every $\varphi(t)$ model is full rank:									
	1053	470	76	191	103	47	86	63	37
# of histories indicates how many occupancy-histories have an expected value greater than one.									

5.10 Discussion

This chapter has dealt with occupancy studies which are conducted over multiple surveys in multiple seasons, as well as possibly being observed between different groups. Occupancy modelling is being used more frequently with better detection technology now available to us. We can see its importance by viewing the high number of papers in the last decade that have used occupancy modelling as a basis for analysis. As these models are being more widely used, obtaining the parameter redundancy of these models is now of interest. We have described the formulation of these models and observed that the occupancy models in this chapter are not intrinsically parameter redundant given a perfect data set, if each season has at least two surveys. Any parameter redundancy that an occupancy model might have (given two or more surveys per season) is due to the sparseness of the observed data. We have explored extrinsic parameter redundancy for two data sets in this chapter, one on house finches in Table 5.1 from MacKenzie et al. (2006) and one on amphibian monitoring in Greater Yellowstone and Grand Teton national parks in Tables 5.2 and 5.3 from Gould et al. (2012). These results convey the importance of rich data sets to enable all of the parameters in the model to be estimated. We further developed `Maple` codes in this chapter to incorporate occupancy-histories with some missing observations during the study which reduces the parameter redundancy of some models. We then evaluated sample size limitations and showed that more occupancy-histories need to be observed when there are low transition and detection probabilities to enable all of the model parameters to be estimated.

We continue examining parameter redundancy in occupancy models in Chapter 6, examining further extensions. This includes the addition of multiple states, a model to consider the interaction of occupancy states between two species, and the Royle-Nichols model which allows animal abundance estimates to be obtained.

Chapter 6

Complex Occupancy Models

In this chapter, we explore a variety of different occupancy models which are extensions of the models explored in Chapter 5. The alternative models include the following:

- We explore a multiple-states occupancy model in Section 6.1 which can account for different states of occupancy detection. For example, this can be to determine whether an animal species is either present and breeding, present and not breeding, or absent, as its possible states of occupancy. We examine how the model can be constructed for three different states to begin with, and then extend further to include observations for different species groups, of which there may be missing observations in the occupancy-histories as well. `Maple` code is given so that we that obtain extrinsic parameter redundancy results and includes a real example involving green frogs in Maryland, taken from MacKenzie et al. (2009).

- We explore a further extension of this multiple-states model where the occupancy state of the species is known before the study is conducted at time $t = 0$. This is explored in Section 6.2 and `Maple` code is provided to examine extrinsic parameter redundancy results.

- We then explore a two-species interaction model in Section 6.3. This model considers two different species simultaneously, allowing the presence/absence of one species to affect the presence/absence of another species. We only explore parameter redundancy in the two-species interaction model but show that a model can be proposed for a multiple number of interacting species. We then examine extrinsic parameter redundancy using the real data example of Sections 5.7 and 5.8, exploring amphibian breeding in Yellowstone and Grand Teton national

parks from Gould et al. (2012).

- We finally explore the Royle-Nichols occupancy model of Royle and Nichols (2003) in Section 6.4. Royle and Nichols (2003) explore the link between animal abundance and species detection to propose an alternative model which takes into account the fact that locations with a higher abundance may have a higher species detection probability. We then examine extrinsic parameter redundancy results from a data set on woodthrush and catbirds from the North America Breeding Bird Survey given in Fiske et al. (2014).

As there is a wide variety of occupancy models explored in this thesis, we show all the occupancy models examined in this thesis in Table 6.1 for clarity. As there is a very large number of possible exhaustive summary terms in the occupancy models explored in this chapter, it is unrealistic that all possible occupancy-histories will be observed for these models. It is for this reason that we *only* explore extrinsic parameter redundancy results and not intrinsic parameter redundancy results in this chapter.

Table 6.1: The different occupancy models examined for parameter redundancy in this thesis

Model	Section	Summary of model
The basic occupancy model	5.2	The basic occupancy model with only one detection survey in a single season
Multiple-surveys occupancy model	5.3	The basic occupancy model with a multiple number of detection surveys during a single season
Multiple-seasons occupancy model	5.4	The basic occupancy model with multiple surveys and multiple seasons during the study
Multiple-groups occupancy model	5.7	The multiple-seasons model with different groups during the study
Multiple-states occupancy model	6.1	The multiple-seasons model with different possible states of detection, e.g. breeding/non-breeding/absence
Multiple-states and multiple-groups model	6.1.3	The multiple-states model which also considers observations at different groups
Known initial state model	6.2	The multiple-states model where an occupancy-history's original state is known
2-species interaction occupancy model	6.3	The multiple-seasons model where we consider an interaction between two different species
The Royle-Nichols model	6.4	An alternative single-season occupancy model where it is possible to obtain estimates of animal abundance.

6.1 The multiple-states occupancy model

This section considers a **multiple-states occupancy model** which is an extension to the occupancy model with the addition of multiple occupied states. A commonly used example of an occupancy study with multiple occupied states is one with three different true occupancy states, where a species could be present and breeding, where a species could be present but is not breeding, and where a species could be absent. This is a sequential model as even if members of a species is detected as non-breeders, we need to consider the possibility that the species was actually breeding but not detected as such.

Multiple-state models are widely used in other ecological models, such as capture-recapture models in Brownie et al. (1993), Nichols and Kendall (1995), Doligez et al. (2002), Choquet et al. (2004) and McCrea et al. (2010, 2012), as well as in capture-recapture-recovery models in McCrea (2012) and King (2012). Parameter redundancy of multiple-state models has previously been examined in Gimenez et al. (2004) and Cole (2012).

Multiple-state occupancy models were developed in MacKenzie et al. (2009), following the methods of Royle and Link (2005) and Nichols et al. (2007) to allow for the changes in a greater number of occupied states. The model given in MacKenzie et al. (2009) is the one we use in this chapter for our parameter redundancy analysis. We first consider the multiple-state occupancy model with only three different states, and then show that the model can be extended to more than three states.

6.1.1 A three-state occupancy model

Consider a model which has three different states which are denoted as being in state 0, state 1 and state 2 respectively. State 0 represents the state where the species is absent from the study location. State 1 and state 2 then represents the states where the species is present, however state 1 and state 2 are different observational true states: For example, state 1 could represent the state where the species is present but not breeding, and state 2 could represent the state where the species is present and breeding. This is a common practical example of a three state occupancy model as given in MacKenzie et al. (2009). Note that if a species is not detected in a season, the species could be in any of the three possible states: Breeding, not breeding, or the species actually was absent at the location. As the observational true states are hierarchical, there are two possibilities if the species is detected, but a species that is

breeding could be wrongly observed as not breeding. We know with certainty that a species is breeding if there is a detection where evidence of breeding is observed.

Let $\varphi_0^{[a]}$ be the initial occupancy probability that the species is in state $a = \{0, 1, 2\}$ at time $t = 0$. Further let $\varphi_t^{[a,b]}$ be the occupancy transition probability of going from state a in season $t-1$ to state b in season t, and observe that $\varphi_t^{[a,0]} = 1 - \varphi_t^{[a,1]} - \varphi_t^{[a,2]}$. We observe repeated surveys during each season where the detection parameters change notation slightly from before so that the probability of species detection is given as $p_{t,j}^{l,m}$ for survey j in season t, where the species is detected at state l given its true state is in state m. We change notation here as we now need to consider a multiple number of observed states while we did not need to do this in Chapter 5. Similarly to the transition probabilities, $p_{t,j}^{0,2} = 1 - p_{t,j}^{1,2} - p_{t,j}^{2,2}$ and $p_{t,j}^{0,1} = 1 - p_{t,j}^{1,1}$ apply here to reduce the number of parameters in the model. We display the series of detections in an occupancy-history to indicate the observed state during each survey in the study. This is similar to before in Chapter 6 but we now obtain a '2' within a species occupancy-history if the species was observed in state 2 during that survey. Examples 6.1 to 6.3 demonstrate how the probabilities of some multiple-state occupancy-histories can be expressed.

Example 6.1: Consider the single-season occupancy-history $h = 1220$. This indicates the species is breeding for the season, but the species was detected as not breeding in the first survey, detected as breeding in the second and third surveys and the species was undetected in the final survey. The species must be in a breeding state as there is at least one detection in state 2 during the season. The probability of this occupancy-history is given as

$$Pr(h) = \varphi_0^{[2]} p_{1,1}^{1,2} p_{1,2}^{2,2} p_{1,3}^{2,2} (1 - p_{1,4}^{1,2} - p_{1,4}^{2,2}).$$

□

Example 6.2: Consider the single-season occupancy-history $h = 1110$. This could mean that the species was breeding but breeding was not detected in the first three surveys and the species was not detected in the final survey, or the species was actually not breeding and the species was detected as such in the first three surveys and was not detected in the final survey. The probability of this occupancy-history is then equal to

$$Pr(h) = \varphi_0^{[2]} p_{1,1}^{1,2} p_{1,2}^{1,2} p_{1,3}^{1,2} (1 - p_{1,4}^{1,2} - p_{1,4}^{2,2}) + \varphi_0^{[1]} p_{1,1}^{1,1} p_{1,2}^{1,1} p_{1,3}^{1,1} (1 - p_{1,4}^{1,1}).$$

Observe the consideration of both possibilities where the species could have been a

breeder but only detected as a non-breeder, or the species truly was a non-breeder. □

Example 6.3: Further consider the single-season occupancy-history if $h = 0000$ with no species detection at all. This could mean the species was present and breeding but not detected in any of the surveys, or the species was present and not breeding but was not detected in any of the surveys, or the species truly was absent at that location. The probability of this occupancy-history is then equal to

$$Pr(h) = \varphi_0^{[2]} \prod_{j=1}^{4} \left(1 - p_{1,j}^{1,2} - p_{1,j}^{2,2}\right) + \varphi_0^{[1]} \prod_{j=1}^{4} \left(1 - p_{1,j}^{1,1}\right) + \left(1 - \varphi_0^{[1]} - \varphi_0^{[2]}\right).$$

This complex expression stems from the fact that we are uncertain about its presence or absence at the site, as well as about its breeding state if the species truly was present at the site but not detected. □

Similarly to Section 5.4 we use matrix notation to express the general probability for an occupancy-history. We extend ϕ_t to be

$$\phi_t = \begin{bmatrix} \varphi_t^{[2,2]} & \varphi_t^{[2,1]} & 1 - \varphi_t^{[2,1]} - \varphi_t^{[2,2]} \\ \varphi_t^{[1,2]} & \varphi_t^{[1,1]} & 1 - \varphi_t^{[1,1]} - \varphi_t^{[1,2]} \\ \varphi_t^{[0,2]} & \varphi_t^{[0,1]} & 1 - \varphi_t^{[0,1]} - \varphi_t^{[0,2]} \end{bmatrix},$$

for a 3-state model for every season $t = 1, \ldots, T-1$, with ϕ_0 given as

$$\phi_0 = \begin{bmatrix} \varphi_0^{[2]} & \varphi_0^{[1]} & 1 - \varphi_0^{[1]} - \varphi_0^{[2]} \end{bmatrix}.$$

We then consider the column vector $\mathbf{p}_{X,t}$, where the first entry is the probability of observing that occupancy-history given that the species is in the breeding state, the second entry is the probability of observing that occupancy-history given that the species is in the non-breeding state, and the last entry is the probability of observing that occupancy-history given that the species is absent, for all seasons $t = 1, \ldots, T$. If there is at least one detection during the season, the last entry would be equal to zero as it is clear the species is present in that season, while the final entry would be equal to one if the species was not observed during that season. Let $\delta_{t,j}$ denote the species observed state at survey j in season t where $\delta_{t,j} = \{0, 1, 2\}$. Further let

$$\alpha_{t,j} = \begin{cases} 1 & \text{if } \delta_{t,j} = 1 \text{ or } 2 \text{ (i.e. for when the species is detected regardless of state)}, \\ 0 & \text{if } \delta_{t,j} = 0 \text{ (i.e. for when the species is not detected)}, \end{cases}$$

for every survey $j = 1, \ldots, k_t$, where k_t is the number of surveys in season t, and

$$\beta_t^0 = \begin{cases} 1 & \text{when } \max(\delta_{t,j}) = 0 \text{ for all } j = 1, \ldots, k_t, \\ 0 & \text{when } \max(\delta_{t,j}) = 1 \text{ or } 2 \text{ for all } j = 1, \ldots, k_t, \end{cases}$$

$$\beta_t^1 = \begin{cases} 1 & \text{when } \max(\delta_{t,j}) = 0 \text{ or } 1 \text{ for all } j = 1, \ldots, k_t, \\ 0 & \text{when } \max(\delta_{t,j}) = 2 \text{ for all } j = 1, \ldots, k_t, \end{cases}$$

for seasons $t = 1, \ldots, T$. The column vector $\mathbf{p}_{X,t}$ is given as

$$\mathbf{p}_{X,t} = \begin{bmatrix} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j} p_{t,j}^{\delta_{t,j},2} + (1 - \alpha_{t,j}) \left(1 - p_{t,j}^{1,2} - p_{t,j}^{2,2}\right) \right\} \\ \beta_t^1 \prod_{j=1}^{k_t} \left\{ \alpha_{t,j} p_{t,j}^{1,1} + (1 - \alpha_{t,j}) \left(1 - p_{t,j}^{1,1}\right) \right\} \\ \beta_t^0 \end{bmatrix}. \tag{6.1}$$

We show some examples of $\mathbf{p}_{X,t}$ vectors for the previous examples 6.1 to 6.3 below.

Example 6.1 revisited: In the example where the occupancy-history is observed as $h = 1220$, the column vector $\mathbf{p}_{1220,t}$ is given as

$$\mathbf{p}_{1220,t} = \begin{bmatrix} p_{t,1}^{1,2} p_{t,2}^{2,2} p_{t,3}^{2,2} (1 - p_{t,4}^{1,2} - p_{t,4}^{2,2}) \\ 0 \\ 0 \end{bmatrix}.$$

□

Example 6.2 revisited: In the example where the occupancy-history is observed as $h = 1110$, the column vector $\mathbf{p}_{1110,t}$ is given as

$$\mathbf{p}_{1110,t} = \begin{bmatrix} p_{t,1}^{1,2} p_{t,2}^{1,2} p_{t,3}^{1,2} (1 - p_{t,4}^{1,2} - p_{t,4}^{2,2}) \\ p_{t,1}^{1,1} p_{t,2}^{1,1} p_{t,3}^{1,1} (1 - p_{t,4}^{1,1}) \\ 0 \end{bmatrix}.$$

□

Example 6.3 revisited: In the final example where the occupancy-history is observed as $h = 0000$ with no detections at all during the season, the column vector $\mathbf{p}_{0000,t}$ is given as

$$\mathbf{p}_{0000,t} = \begin{bmatrix} (1 - p_{t,1}^{1,2} - p_{t,1}^{2,2})(1 - p_{t,2}^{1,2} - p_{t,2}^{2,2})(1 - p_{t,3}^{1,2} - p_{t,3}^{2,2})(1 - p_{t,4}^{1,2} - p_{t,4}^{2,2}) \\ (1 - p_{t,1}^{1,1})(1 - p_{t,2}^{1,1})(1 - p_{t,3}^{1,1})(1 - p_{t,4}^{1,1}) \\ 1 \end{bmatrix}.$$

□

Observe that while this formulation of the model given in Equation (6.1) is similar to the form given in MacKenzie et al. (2009), we have switched the order of states in $\mathbf{p}_{X,t}$ so that the first entry denotes the probability the species is in state 2 rather than denoting the species is in state 0 as in MacKenzie et al. (2009). We have formulated the model this way as we feel the formulation is easier to show in examples if the first entry of $\mathbf{p}_{X,t}$ is the probability of being in state 2, rather than in state 0.

Both the formulation here and in MacKenzie et al. (2009) give the same probability for any occupancy-history, as the transition matrices are then transposed in MacKenzie et al. (2009) to accommodate for the alternative expression for $\mathbf{p}_{X,t}$.

Let $\Delta(\mathbf{p}_{X,t})$ be the diagonal matrix with the elements of $\mathbf{p}_{X,t}$ along the main diagonal in order and zeros for the non-diagonal entries. The probability of a certain occupancy-history can be expressed as

$$Pr(h_X) = \phi_0 \left[\prod_{t=1}^{T-1} \Delta(\mathbf{p}_{X,t}) \phi_t \right] \mathbf{p}_{X,T}.$$

The likelihood is then given as

$$L = \prod_{X=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study. We give Examples 6.4 and 6.5 to show how the probabilities of occupancy-histories for a multiple number of seasons can be obtained.

Example 6.4: Consider a study where there are three seasons where each season contains two surveys, and three different observational states are considered. The probability of the occupancy-history $h = 21|22|11$ can be found by the product of the

following terms

$$
\begin{aligned}
Pr(h) &= \phi_0 \Delta(\mathbf{p}_{21,1}) \phi_1 \Delta(\mathbf{p}_{22,2}) \phi_2 \mathbf{p}_{11,3} \\
&= \begin{bmatrix} \varphi_0^{[2]} & \varphi_0^{[1]} & 1-\varphi_0^{[1]}-\varphi_0^{[2]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}^{2,2} p_{1,2}^{1,2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} \varphi_1^{[2,2]} & \varphi_1^{[2,1]} & 1-\varphi_1^{[2,1]}-\varphi_1^{[2,2]} \\ \varphi_1^{[1,2]} & \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]}-\varphi_1^{[1,2]} \\ \varphi_1^{[0,2]} & \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]}-\varphi_1^{[0,2]} \end{bmatrix} \cdot \begin{bmatrix} p_{2,1}^{2,2} p_{2,2}^{2,2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} \varphi_2^{[2,2]} & \varphi_2^{[2,1]} & 1-\varphi_2^{[2,1]}-\varphi_2^{[2,2]} \\ \varphi_2^{[1,2]} & \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]}-\varphi_2^{[1,2]} \\ \varphi_2^{[0,2]} & \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]}-\varphi_2^{[0,2]} \end{bmatrix} \cdot \begin{bmatrix} p_{3,1}^{1,2} p_{3,2}^{1,2} \\ p_{3,1}^{1,1} p_{3,2}^{1,1} \\ 0 \end{bmatrix} \\
&= \varphi_0^{[2]} p_{1,1}^{[2,2]} p_{1,2}^{[1,2]} \varphi_1^{[2,2]} p_{2,1}^{[2,2]} p_{2,2}^{[2,2]} \left(\varphi_2^{[2,2]} p_{3,1}^{[1,2]} p_{3,2}^{[1,2]} + \varphi_2^{[2,1]} p_{3,1}^{[1,1]} p_{3,2}^{[1,1]} \right).
\end{aligned}
$$

This final probability can be expanded into two cases, where the species could be in a breeding state throughout the study, or the species could be in a breeding state for the first two seasons and be in a non-breeding state in the final season. □

Example 6.5: Consider the same study where the occupancy-history $h = 10|00|02$ is observed. The probability of this occupancy-history can be expressed by the product

of the terms

$$
\begin{aligned}
Pr(h) &= \phi_0\Delta(\mathbf{p}_{10,1})\phi_1\Delta(\mathbf{p}_{00,2})\phi_2\mathbf{p}_{02,3} \\
&= \begin{bmatrix} \varphi_0^{[2]} & \varphi_0^{[1]} & 1-\varphi_0^{[1]}-\varphi_0^{[2]} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1}^{1,2}(1-p_{1,2}^{1,2}-p_{1,2}^{2,2}) & 0 & 0 \\ 0 & p_{1,1}^{1,1}(1-p_{1,2}^{1,1}) & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} \varphi_1^{[2,2]} & \varphi_1^{[2,1]} & 1-\varphi_1^{[2,1]}-\varphi_1^{[2,2]} \\ \varphi_1^{[1,2]} & \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]}-\varphi_1^{[1,2]} \\ \varphi_1^{[0,2]} & \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]}-\varphi_1^{[0,2]} \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} (1-p_{2,1}^{1,2}-p_{2,1}^{2,2})(1-p_{2,2}^{1,2}-p_{2,2}^{2,2}) & 0 & 0 \\ 0 & (1-p_{2,1}^{1,1})(1-p_{2,2}^{1,1}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} \varphi_2^{[2,2]} & \varphi_2^{[2,1]} & 1-\varphi_2^{[2,1]}-\varphi_2^{[2,2]} \\ \varphi_2^{[1,2]} & \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]}-\varphi_2^{[1,2]} \\ \varphi_2^{[0,2]} & \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]}-\varphi_2^{[0,2]} \end{bmatrix} \cdot \begin{bmatrix} (1-p_{3,1}^{1,2}-p_{3,1}^{2,2})p_{3,2}^{2,2} \\ 0 \\ 0 \end{bmatrix} \\
&= \left\{ \left[\varphi_0^{[2]} p_{1,1}^{1,2} \left(1-p_{1,2}^{1,2}-p_{1,2}^{2,2}\right) \varphi_1^{[2,2]} + \varphi_0^{[1]} p_{1,1}^{1,1} \left(1-p_{1,2}^{1,1}\right) \varphi_1^{[1,2]} \right] \cdot \right. \\
&\quad \left(1-p_{2,1}^{1,2}-p_{2,1}^{2,2}\right) \left(1-p_{2,2}^{1,2}-p_{2,2}^{2,2}\right) \varphi_2^{[2,2]} + \left[\varphi_0^{[2]} p_{1,1}^{1,2} \left(1-p_{1,2}^{1,2}-p_{1,2}^{2,2}\right) \varphi_1^{[2,1]} \right. \\
&\quad \left. + \varphi_0^{[1]} p_{1,1}^{1,1} \left(1-p_{1,2}^{1,1}\right) \varphi_1^{[1,1]} \right] \cdot \left(1-p_{2,1}^{1,1}\right) \left(1-p_{2,2}^{1,1}\right) \varphi_2^{[1,2]} \\
&\quad + \left[\varphi_0^{[2]} p_{1,1}^{1,2} \left(1-p_{1,2}^{1,2}-p_{1,2}^{2,2}\right) \left(1-\varphi_1^{[2,1]}-\varphi_1^{[2,2]}\right) \varphi_2^{[0,2]} \right. \\
&\quad \left. \left. + \varphi_0^{[1]} p_{1,1}^{1,1} \left(1-p_{1,2}^{1,1}\right) \left(1-\varphi_1^{[1,1]}-\varphi_1^{[1,2]}\right) \varphi_2^{[0,2]} \right] \right\} \left(1-p_{3,1}^{1,2}-p_{3,1}^{2,2}\right) p_{3,2}^{2,2}.
\end{aligned}
$$

This rather complex probability is a consequence of being unsure if the species is present, either as breeding or non-breeding, or absent in the second season. Complication in occupancy models arise when there are no detections in a single season, as shown previously in Example 5.4 of Section 5.4. □

6.1.2 A general m-state model which involves group-dependence

We now consider a more general model for m different states for s different groups. This model can include a range of different parameter dependencies including group-dependency. It is important to note that the ambiguity of the observations is only sequential, i.e. if a species is detected in state 1, while the species may actually be in a state higher than 1, it is clear the species is *not* in state 0. This means we have the following cell probabilities as given in Table 6.2. As is shown in Table 6.2, if the

Table 6.2: The probability of a species true state given its observed state

True State	Observed State					
	0	1	2	$\cdots$	$l-1$	l
0	1	0	0	$\cdots$	0	0
1	$1-p_{t,j,s}^{1,1}$	$p_{t,j,s}^{1,1}$	0	$\cdots$	0	0
2	$1-p_{t,j,s}^{1,2}-p_{t,j,s}^{2,2}$	$p_{t,j,s}^{1,2}$	$p_{t,j,s}^{2,2}$	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$m-1$	$1-\left(\sum_{l=1}^{m-1} p_{t,j,s}^{l,m-1}\right)$	$p_{t,j,s}^{1,m-1}$	$p_{t,j,s}^{2,m-1}$	$\cdots$	$p_{t,j,s}^{l-1,m-1}$	0
m	$1-\left(\sum_{l=1}^{m} p_{t,j,s}^{l,m}\right)$	$p_{t,j,s}^{1,m}$	$p_{t,j,s}^{2,m}$	$\cdots$	$p_{t,j,s}^{l-1,m}$	$p_{t,j,s}^{l,m}$

species is truly in state m, then the species could be observed at any state in the study, though it is hoped to be observed in state m.

The transition matrix is now given by

$$\phi_{t,s} = \begin{bmatrix} \varphi_{t,s}^{[m,m]} & \varphi_{t,s}^{[m,m-1]} & \cdots & \varphi_{t,s}^{[m,1]} & 1-\sum_{l=1}^{m}\varphi_{t,s}^{[m,l]} \\ \varphi_{t,s}^{[m-1,m]} & \varphi_{t,s}^{[m-1,m-1]} & \cdots & \varphi_{t,s}^{[m-1,1]} & 1-\sum_{l=1}^{m}\varphi_{t,s}^{[m-1,l]} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \varphi_{t,s}^{[1,m]} & \varphi_{t,s}^{[1,m-1]} & \cdots & \varphi_{t,s}^{[1,1]} & 1-\sum_{l=1}^{m}\varphi_{t,s}^{[1,l]} \\ \varphi_{t,s}^{[0,m]} & \varphi_{t,s}^{[0,m-1]} & \cdots & \varphi_{t,s}^{[0,1]} & 1-\sum_{l=1}^{m}\varphi_{t,s}^{[0,l]} \end{bmatrix},$$

for every season $t = 1, \ldots, T-1$, with $\phi_{0,s}$ given as

$$\phi_{0,s} = \begin{bmatrix} \varphi_{0,s}^{[m]} & \varphi_{0,s}^{[m-1]} & \cdots & \varphi_{0,s}^{[2]} & \varphi_{0,s}^{[1]} & 1-\sum_{l=1}^{m}\varphi_{0,s}^{[l]} \end{bmatrix}.$$

If we let the observed state in season t at survey j for group s be equal to $\delta_{t,j,s}$, we can define $\alpha_{t,j,s}$ as

$$\alpha_{t,j,s} = \begin{cases} 1 & \text{when } \delta_{t,j,s} \geq 1 \text{ (i.e. for when the species is detected),} \\ 0 & \text{when } \delta_{t,j,s} = 0 \text{ (i.e. for when the species is not detected),} \end{cases}$$

for every survey $j = 1, \ldots, k_t$, and

$$\beta_{t,s}^{l} = \begin{cases} 1 & \text{when } \max(\delta_{t,j,s}) \leq l \text{ for all } j = 1, \ldots, k_t, \\ 0 & \text{when } \max(\delta_{t,j,s}) > l \text{ for all } j = 1, \ldots, k_t, \end{cases}$$

in season $t = 1, \ldots, T$ for group $s = 1, \ldots S$ for every state $l = 0, \ldots, m-1$. The

detection probabilities in the column vector $\mathbf{p}_{X,t,s}$ can be expressed as

$$\mathbf{p}_{X,t,s} = \begin{bmatrix} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m} + (1-\alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m} p_{t,j,s}^{l,m}\right) \right\} \\ \beta_{t,s}^{m-1} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m-1} + (1-\alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m-1} p_{t,j,s}^{l,m-1}\right) \right\} \\ \beta_{t,s}^{m-2} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m-2} + (1-\alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m-2} \beta_{t,j,s}^{l,m-2}\right) \right\} \\ \vdots \\ \beta_{t,s}^{2} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},2} + (1-\alpha_{t,j,s}) \left(1 - p_{t,j,s}^{1,2} - p_{t,j,s}^{2,2}\right) \right\} \\ \beta_{t,s}^{1} \prod_{j=1}^{k_t} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},1} + (1-\alpha_{t,j,s}) \left(1 - p_{t,j,s}^{1,1}\right) \right\} \\ \beta_{t,s}^{0} \end{bmatrix}. \quad (6.2)$$

The probability of an individual occupancy-history is given by

$$Pr(h_X) = \phi_{0,s} \left[\prod_{t=1}^{T-1} \Delta(\mathbf{p}_{X,t,s}) \phi_{t,s} \right] \mathbf{p}_{X,T,s},$$

and the final likelihood can be expressed as

$$L = \prod_{X=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study. We give Example 6.6 to show how the probabilities of occupancy-histories for a large number of multiple seasons can be obtained.

Example 6.6: Consider a study where there are two seasons where each season contains two surveys, along with the species possibly being in four different observational states. The probability of the occupancy-history $h = 42|01$ can be found by the product of

the following terms

$$
\begin{aligned}
Pr(h) &= \phi_0 \Delta(\mathbf{p}_{42,1}) \phi_1 \mathbf{p}_{01,2} \\
&= \left[\begin{array}{ccccc} \varphi_0^{[4]} & \varphi_0^{[3]} & \varphi_0^{[2]} & \varphi_0^{[1]} & 1 - \sum_{l=1}^{4} \varphi_0^{[l]} \end{array} \right] \cdot \left[\begin{array}{ccccc} p_{1,1}^{4,4} p_{1,2}^{2,4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \cdot \\
&\quad \left[\begin{array}{ccccc} \varphi_1^{[4,4]} & \varphi_1^{[4,3]} & \varphi_1^{[4,2]} & \varphi_1^{[4,1]} & 1 - \varphi_1^{[4,1]} - \varphi_1^{[4,2]} - \varphi_1^{[4,3]} - \varphi_1^{[4,4]} \\ \varphi_1^{[3,4]} & \varphi_1^{[3,3]} & \varphi_1^{[3,2]} & \varphi_1^{[3,1]} & 1 - \varphi_1^{[3,1]} - \varphi_1^{[3,2]} - \varphi_1^{[3,3]} - \varphi_1^{[3,4]} \\ \varphi_1^{[2,4]} & \varphi_1^{[2,3]} & \varphi_1^{[2,2]} & \varphi_1^{[2,1]} & 1 - \varphi_1^{[2,1]} - \varphi_1^{[2,2]} - \varphi_1^{[2,3]} - \varphi_1^{[2,4]} \\ \varphi_1^{[1,4]} & \varphi_1^{[1,3]} & \varphi_1^{[1,2]} & \varphi_1^{[1,1]} & 1 - \varphi_1^{[1,1]} - \varphi_1^{[1,2]} - \varphi_1^{[1,3]} - \varphi_1^{[1,4]} \\ \varphi_1^{[0,4]} & \varphi_1^{[0,3]} & \varphi_1^{[0,2]} & \varphi_1^{[0,1]} & 1 - \varphi_1^{[0,1]} - \varphi_1^{[0,2]} - \varphi_1^{[0,3]} - \varphi_1^{[0,4]} \end{array} \right] \cdot \\
&\quad \left[\begin{array}{c} (1 - p_{3,1}^{1,4} - p_{3,1}^{2,4} - p_{3,1}^{3,4} - p_{3,1}^{4,4}) p_{3,2}^{1,4} \\ (1 - p_{3,1}^{1,3} - p_{3,1}^{2,3} - p_{3,1}^{3,3}) p_{3,2}^{1,3} \\ (1 - p_{3,1}^{1,2} - p_{3,1}^{2,2}) p_{3,2}^{1,2} \\ (1 - p_{3,1}^{1,1}) p_{3,2}^{1,1} \\ 0 \end{array} \right] \\
&= \varphi_0^{[4]} p_{1,1}^{4,4} p_{1,2}^{2,4} \left\{ \varphi_1^{[4,4]} \left(1 - p_{3,1}^{1,4} - p_{3,1}^{2,4} - p_{3,1}^{3,4} - p_{3,1}^{4,4} \right) p_{3,2}^{1,4} \right. \\
&\quad + \varphi_1^{[4,3]} \left(1 - p_{3,1}^{1,3} - p_{3,1}^{2,3} - p_{3,1}^{3,3} \right) p_{3,2}^{1,3} + \varphi_1^{[4,2]} \left(1 - p_{3,1}^{1,2} - p_{3,1}^{2,2} \right) p_{3,2}^{1,2} \\
&\quad \left. + \varphi_1^{[4,1]} \left(1 - p_{3,1}^{1,1} \right) p_{3,2}^{1,1} \right\}.
\end{aligned}
$$

□

6.1.3 A general group-dependent m-state model which contains missing observations

We can again account for missing observations in occupancy-histories as shown previously in Section 5.8. In a similar way to before, the way to deal with missing observations is to effectively ignore any parameters relating to that detection survey. As we previously denoted in Section 5.8, we use an underline '_' when viewing an individual occupancy-history where that survey has a missing observation, and we use the notation '−1' in the `Maple` input to denote when that survey has a missing observation. If we continue to let the observed state in an occupancy-history in survey

j in season t for group s be denoted as $\delta_{t,j,s}$, and further specify $\gamma_{t,j,s}$ as

$$\gamma_{t,j,s} = \begin{cases} 1 & \text{when } \delta_{t,j,s} \geq 0 \text{ (i.e. it not a missing observation),} \\ 0 & \text{when } \delta_{t,j,s} = -1 \text{ (i.e. it } \textit{is} \text{ a missing observation).} \end{cases}$$

Let

$$\alpha_{t,j,s} = \begin{cases} 1 & \text{when } \delta_{t,j,s} \geq 1 \text{ (i.e. for when the species is detected regardless of state),} \\ 0 & \text{when } \delta_{t,j,s} = 0 \text{ or } -1 \text{ (i.e. for when the species is not detected),} \end{cases}$$

for every survey $j = 1, \ldots, k_t$ and

$$\beta^l_{t,s} = \begin{cases} 1 & \text{when } \max(\delta_{t,j,s}) \leq l \text{ for all } j = 1, \ldots, k_t, \\ 0 & \text{when } \max(\delta_{t,j,s}) > l \text{ for all } j = 1, \ldots, k_t, \end{cases}$$

in season $t = 1, \ldots, T$ for group $s = 1, \ldots S$. We then define $\mathbf{p}_{X,t,s}$ to be the column vector

$$\mathbf{p}_{X,t,s} = \begin{bmatrix} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m} + (1 - \alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m} p_{t,j,s}^{l,m} \right) \right\} + (1 - \gamma_{t,j,s}) \\ \beta_{t,s}^{m-1} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m-1} + (1 - \alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m-1} p_{t,j,s}^{l,m-1} \right) \right\} + (1 - \gamma_{t,j,s}) \\ \beta_{t,s}^{m-2} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},m-2} + (1 - \alpha_{t,j,s}) \left(1 - \sum_{l=1}^{m-2} p_{t,j,s}^{l,m-2} \right) \right\} + (1 - \gamma_{t,j,s}) \\ \vdots \\ \beta_{t,s}^{2} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},2} + (1 - \alpha_{t,j,s}) \left(1 - p_{t,j,s}^{1,2} - p_{t,j,s}^{2,2} \right) \right\} + (1 - \gamma_{t,j,s}) \\ \beta_{t,s}^{1} \prod_{j=1}^{k_t} \gamma_{t,j,s} \left\{ \alpha_{t,j,s} p_{t,j,s}^{\delta_{t,j,s},1} + (1 - \alpha_{t,j,s}) \left(1 - p_{t,j,s}^{1,1} \right) \right\} + (1 - \gamma_{t,j,s}) \\ \beta_{t,s}^{0} \end{bmatrix}. \tag{6.3}$$

This means as previously that the probability of an individual occupancy-history is equal to

$$Pr(h_X) = \phi_{0,s} \left[\prod_{t=1}^{T-1} \Delta(\mathbf{p}_{X,t,s}) \phi_{t,s} \right] \mathbf{p}_{X,T,s}, \tag{6.4}$$

making the likelihood equal to

$$L = \prod_{X=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study. Example 6.7 below shows how missing observations can be incorporated into this multiple-state model to obtain an occupancy-history probability.

Example 6.7: Consider a study where there are three seasons where each season contains two surveys, along with the species possibly being in three different observational states. The probability of the occupancy-history $h = 1\_|\_|20$ can be found by the product of the following terms

$$\begin{aligned}
Pr(h_7) &= \phi_0\Delta(\mathbf{p}_{1_,1})\phi_1\Delta(\mathbf{p}_{_,2})\phi_2\mathbf{p}_{00,3} \\
&= \left[\begin{array}{ccc} \varphi_0^{[2]} & \varphi_0^{[1]} & 1-\varphi_0^{[1]}-\varphi_0^{[2]} \end{array}\right] \cdot \left[\begin{array}{ccc} p_{1,1}^{1,2} & 0 & 0 \\ 0 & p_{1,1}^{1,1} & 0 \\ 0 & 0 & 0 \end{array}\right] \cdot \\
&\quad \left[\begin{array}{ccc} \varphi_1^{[2,2]} & \varphi_1^{[2,1]} & 1-\varphi_1^{[2,1]}-\varphi_1^{[2,2]} \\ \varphi_1^{[1,2]} & \varphi_1^{[1,1]} & 1-\varphi_1^{[1,1]}-\varphi_1^{[1,2]} \\ \varphi_1^{[0,2]} & \varphi_1^{[0,1]} & 1-\varphi_1^{[0,1]}-\varphi_1^{[0,2]} \end{array}\right] \cdot \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \cdot \\
&\quad \left[\begin{array}{ccc} \varphi_2^{[2,2]} & \varphi_2^{[2,1]} & 1-\varphi_2^{[2,1]}-\varphi_2^{[2,2]} \\ \varphi_2^{[1,2]} & \varphi_2^{[1,1]} & 1-\varphi_2^{[1,1]}-\varphi_2^{[1,2]} \\ \varphi_2^{[0,2]} & \varphi_2^{[0,1]} & 1-\varphi_2^{[0,1]}-\varphi_2^{[0,2]} \end{array}\right] \cdot \left[\begin{array}{c} p_{3,1}^{2,2}(1-p_{3,2}^{2,2}-p_{3,2}^{1,2}) \\ 0 \\ 0 \end{array}\right] \\
&= \left[\varphi_0^{[2]}p_{1,1}^{1,2}\left\{\varphi_1^{[2,2]}\varphi_2^{[2,2]}+\varphi_1^{[2,1]}\varphi_2^{[1,2]}+\left(1-\varphi_1^{[2,2]}-\varphi_1^{[2,1]}\right)\varphi_2^{[0,2]}\right\}+\right. \\
&\quad \left.\varphi_0^{[1]}p_{1,1}^{1,1}\left\{\varphi_1^{[1,2]}\varphi_2^{[2,2]}+\varphi_1^{[1,1]}\varphi_2^{[1,2]}+\left(1-\varphi_1^{[1,2]}-\varphi_1^{[1,1]}\right)\varphi_2^{[0,2]}\right\}\right] \\
&\quad \times p_{3,1}^{2,2}\left(1-p_{3,2}^{2,2}-p_{3,2}^{1,2}\right).
\end{aligned}$$

□

6.1.4 Extrinsic parameter redundancy and Maple code for the multiple-states model

The `Maple` procedures `occmodstateshist` and `occmodstategroupshist` given in the electronic appendices `example6.8.mw` and `example6.9.mw` are extensions of the code from Section 5.8 where we are now considering more than two states in the model. The procedure `occmodstateshist` given in Example 6.8 is for a multiple-state model without any group-dependency and the procedure `occmodstategroupshist` given in Example 6.9 is for a multiple-state model *with* possible group-dependent parameters. Note that if only two states are listed in the code's input, this code generates essentially the same output as the code for the multiple-seasons model of the previous chapter as they give identical exhaustive summaries but with different parameter notations. This multiple-states model is flexible as you can relax similar assumptions to previously in this model: The transition probabilities can be constant, season-dependent and/or

state-dependent, and the detection probabilities can be constant, season-dependent and/or state-dependent and/or survey-dependent, i.e. there are 32 different possibilities of model parameters. The second procedure in Example 6.9 can also have parameters which are group-dependent as well.

Example 6.8 - The multiple-states occupancy model: [See electronic appendix `example6.8.mw`] Take an example of these occupancy-histories for a three-state model for two seasons with two surveys per season: 22|21, 22|20, 21|22, 21|01, 21|00, 20|12, 20|01, 20|00, 12|22, 12|12, 12|21, 12|02, 12|00, 02|21, 02|12, 02|01, 01|22, 01|12, 01|01, 00|22, 00|20, 00|10, 00|02, 00|01 and 00|00. The first input in the `Maple` procedure `occmodstateshist` is a row vector of how many surveys there are per season, given in the vector $\mathbf{S}$ as

$$\mathbf{S} = \begin{bmatrix} 2 & 2 \end{bmatrix}.$$

The second input is the observed data of the occupancy-histories given in the matrix $\mathbf{T}$ as

$$\mathbf{T} = \begin{bmatrix} 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 0 \\ 2 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 \\ 2 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Consider a model where all the parameters are season- and state-dependent, along with survey-dependence for the detection parameters. The parameter redundancy of that occupancy model can be found by the code below:

```
S := < <2|2> >;
T := < <2|2|2|1>, <2|2|2|0>, <2|1|2|2>, <2|1|0|1>, <2|1|0|0>,
<2|0|1|2>, <2|0|1|2>, <2|0|0|1>, <2|0|0|0>, <1|2|2|2>, <1|2|1|2>,
<1|2|2|1>, <1|2|0|2>, <1|2|0|0>, <0|2|2|1>, <0|2|1|2>, <0|2|0|1>,
<0|1|2|2>, <0|1|2|2>, <0|1|1|2>, <0|1|0|1>, <0|0|2|2>, <0|0|2|0>,
<0|0|1|0>, <0|0|0|2>, <0|0|0|1>, <0|0|0|0> >;
kappa := occmodstateshist(3,4,8,S,T);
# Inputs: (M,y,z,S,T);
# M = number of states;
# y = transition probability; z = detection probability;
# for y: '1'=constant, '2'=season-dep, '3'=state-dep,
# '4'=season+state-dep;
```

```
# for z: '1'=constant, '2'=season-dep, '3'=survey-dep,
# '4'=season+survey-dep, '5'=state-dep,'6'=season+state-dep,
# '7'=survey+state-dep, '8'=season+survey+state-dep;
# S = row vector of the number of surveys per season;
# T = list of all occupancy-histories.
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

This generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_{0,2}p_{1,1,2,2}p_{1,2,2,2}\varphi_{1,2,2}p_{2,1,2,2}p_{2,2,1,2} \\ \varphi_{0,2}p_{1,1,2,2}p_{1,2,2,2}\varphi_{1,2,2}p_{2,1,2,2}(1-p_{2,2,1,2}-p_{2,2,2,2}) \\ \varphi_{0,2}p_{1,1,2,2}p_{1,2,1,2}\varphi_{1,2,2}p_{2,1,2,2}p_{2,2,2,2} \\ \varphi_{0,2}p_{1,1,2,2}p_{1,2,1,2}\left\{\varphi_{1,2,2}(1-p_{2,1,1,2}-p_{2,1,2,2})p_{2,2,1,2}+\varphi_{1,2,1}(1-p_{2,1,1,1})p_{2,2,1,1}\right\} \\ \varphi_{0,2}p_{1,1,2,2}p_{1,2,1,2}\left\{\varphi_{1,2,2}(1-p_{2,1,1,2}-p_{2,1,2,2})(1-p_{2,2,1,2}-p_{2,2,2,2})+\cdots\right. \\ \left.\cdots\varphi_{1,2,1}(1-p_{2,1,1,1})(1-p_{2,2,1,1})+(1-\varphi_{1,2,1}-\varphi_{1,2,2})\right\} \\ \vdots \end{bmatrix},$$

where $\varphi_{0,l} = \varphi_0^{[l]}$, $\varphi_{t,l,m} = \varphi_t^{[l,m]}$ and $p_{t,j,l,m} = p_{t,j}^{l,m}$ for survey j in season t with observed state l at true state m. `Maple` does not perform operations well for parameters which have superscripts, so we have used a number of subscripts instead to denote different parameters instead. The parameter set and derivative matrix are not given here to save space but can be found in the `Maple` file `example6.8.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{17}, \mathbf{d} := \mathbf{3}$.

This exemplar data set results in a parameter deficiency of 3 with 17 estimable parameter combinations. This result can then be contrasted with those for different models, for instance if the transition probabilities were relaxed to be only season-dependent and the detection probabilities relaxed to be only season- and state-dependent, there is then no parameter redundancy in the model. □

Next we consider a group-dependent model, where we demonstrate how the `Maple` procedure `occmodstategroupshist` can be used to generate exhaustive summary terms

in Example 6.9.

Example 6.9 - The multiple-states occupancy model with group-dependency: [See electronic appendix `example6.9.mw`] Consider a three-state, two-season study with two surveys per season, where we take the 25 occupancy-histories from Example 6.8 as being in the first group. We now observe 17 different occupancy-histories from a second group: 22|21, 21|01, 20|12, 20|02, 20|01, 20|00, 12|22, 12|12, 12|21, 12|00, 02|21, 02|12, 02|01, 00|20, 00|10, 00|02 and 00|00. The first input in the `Maple` procedure `occmodstategroupshist` is a row vector of how many surveys there are per season, given in the vector **S** as

$$\mathbf{S} = \begin{bmatrix} 2 & 2 \end{bmatrix}.$$

We now have a further input for this procedure unlike the procedure `occmodstateshist` which does not consider group-dependence. We denote **S2** as the column vector denoting which group each occupancy-history is from. This is given as

$$\mathbf{S2} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 2 & 2 & \cdots & 2 \end{bmatrix}^T.$$

The final input is the observed data of the occupancy-histories given in the matrix **T** as

$$\mathbf{T} = \begin{bmatrix} 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 1 \\ 2 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consider a model where all the parameters are season-, state- and group-dependent, along with survey-dependence for the detection parameters. The parameter redundancy of that occupancy model can be found by using the code below:

```
S := < <2|2> >;
S2 := < <1>, ... (25 1's in total) ... <1>,
<2>, ... (17 2's in total) ... <2> >;
T := < <2|2|2|1>, <2|2|2|0>, <2|1|2|2>, <2|1|0|1>, <2|1|0|0>,
```

```
... (all occupancy-histories from group 1) ... <0|0|0|1>, <0|0|0|0>,
<2|2|2|1>, <2|1|0|1>, <2|0|1|2>, <2|0|0|2>, <2|0|0|1>, <2|0|0|0>,
... (all occupancy-histories from group 2) ... <0|0|0|2>, <0|0|0|0> >;
kappa := occmodstategroupshist(3,8,16,S,S2,T);
# Inputs: (M,y,z,S,S2,T);
# M = number of states;
# y = transition probability; z = detection probability;
# for y: '1'=constant, '2'=season-dep, '3'=state-dep,
# '4'=season+state-dep, '5'=group-dep, '6'=season+group-dep,
# '7'=state+group-dep, '8'=season+state+group-dep;
#  for z: '1'=constant, '2'=season-dep, '3'=survey-dep,
# '4'=season+survey-dep, '5'=state-dep, '6'=season+state-dep,
# '7'=survey+state-dep, '8'=season+survey+state-dep,
# '9'=group-dep, '10'=season+group-dep, '11'=survey+group-dep,
# '12'=season+survey+group-dep, '13'=state+group-dep,
# '14'=season+state+group-dep, '15'=survey+state+group-dep,
# '16'=season+survey+state+group-dep;
# S = row vector of the number of surveys per season;
# S2 = column vector determining the group of each occupancy-history;
# T = list of all occupancy-histories.
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

This generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_{0,2,1}p_{1,1,2,2,1}p_{1,2,2,2,1}\varphi_{1,2,2,1}p_{2,1,2,2,1}p_{2,2,1,2,1} \\ \varphi_{0,2,1}p_{1,1,2,2,1}p_{1,2,2,2,1}\varphi_{1,2,2,1}p_{2,1,2,2,1}(1-p_{2,2,1,2,1}-p_{2,2,2,2,1}) \\ \vdots \\ \varphi_{0,2,2}p_{1,1,2,2,2}p_{1,2,2,2,2}\varphi_{1,2,2,2}p_{2,1,2,2,2}p_{2,2,1,2,2} \\ \varphi_{0,2,2}p_{1,1,2,2,2}p_{1,2,1,2,2}\left\{\varphi_{1,2,2,2}(1-p_{2,1,1,2,2}-p_{2,1,2,2,2})p_{2,2,1,2,2}\cdots \right. \\ \left. +\varphi_{1,2,1,2}(1-p_{2,1,1,1,2})p_{2,2,1,1,2}\right\} \end{bmatrix},$$

where $\varphi_{0,l,s} = \varphi_{0,s}^{[l]}$, $\varphi_{t,l,m,s} = \varphi_{t,s}^{[l,m]}$ and $p_{t,j,l,m,s} = p_{t,j,s}^{l,m}$ for survey j in season t with observed state l at true state m. The parameter set and derivative matrix, which are not given here to save space but can be found in the `Maple` file `example6.9.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the

model rank and parameter deficiency as

$\mathbf{r} := \mathbf{31}, \mathbf{d} := \mathbf{9}$.

This exemplar data set results in a parameter deficiency of 9 with 31 estimable parameter combinations. □

6.1.5 A real data example: Green frogs in Maryland

We now illustrate extrinsic parameter redundancy results for multiple-state models, using the `Maple` procedure `occmodstateshist` from Example 6.8, on a data set of green frogs in Maryland from MacKenzie et al. (2009).

Example 6.10 - The multiple-states occupancy model for the green frogs data set: We now consider a data set of green frogs from Maryland, USA, observed as part of the North American Amphibian Monitoring Program (NAAMP). This data set comes from MacKenzie et al. (2009) and consists of 10 listening stations spaced around a 24.1km route where each station is at least 0.8km apart. The green frogs (*Rana clamitans*) were observed in Maryland, USA from 2001 to 2005 inclusive. More details on the NAAMP is given in Weir and Mossman (2005) and this data set has been analysed previously in Royle and Link (2005) and Weir et al. (2005, 2009). A full analysis of parameter estimate results from this data set are given in MacKenzie et al. (2009) and for the 2001 year only data in Royle and Link (2005). 160 out of the 281 different occupancy-histories involve some missing observations where the observations were not recorded.

The observers in the study listen to green frogs at stations along the route and they do so for a total of five minutes at each station. There are four different classification states for the frogs, with the different states detailed as follows:

- A '0' denotes when no green frogs were detected at that survey.
- A '1' denotes the presence of green frogs if some of them can be counted during the survey.
- A '2' denotes when individual frog calls can be distinguished with some overlapping of calls.
- A '3' denotes some constant, overlapping calls when there is a full chorus of frogs heard.

It is assumed that any station is able to achieve any of these four states and it is a sequential model, i.e. if the species is detected in state 1, then the species may actually be in state 2 or state 3, but the species is definitely not in state 0; however if the species is detected in state 3, then the species is truly in state 3.

We use the notation where the transition probabilities can be constant, $\varphi(\cdot)$, season-dependent, $\varphi(t)$, and/or state-dependent, $\varphi(m)$, and where the detection probability can be constant, $p(\cdot)$, season-dependent, $p(t)$, and/or survey-dependent, $p(j)$, and/or state-dependent, $p(m)$. For example, $\varphi(\cdot)p(t, j, m)$ is a model with a constant transition probability with the detection probabilities dependent on season, survey and state. We consider two cases in our analysis, one case where we omit occupancy-histories containing any missing observations, and a second case where we include these partial histories in the analysis. The case where occupancy-histories with missing observations are deleted has 68 distinct occupancy-histories out of the 121 individual occupancy-histories observed, and the case where occupancy-histories with missing observations are included has 178 distinct occupancy-histories out of the 281 individual occupancy-histories observed. We present extrinsic parameter redundancy results in Table 6.3.

Table 6.3 displays the result that the $\varphi(t, m)$ $p(t, j, m)$ model has *less* parameter deficiency than the $\varphi(\cdot)$ $p(t, j, m)$ model, i.e. adding season- and state-dependency to the φ parameters *decreases* the parameter deficiency of the model. The result that increasing model complexity helps decrease parameter redundancy is not generally true, but the results show that a particular confounding of some of the parameters is eliminated when more parameters are added to the model. This is similar to a result found in Section 2.9 where the T/A mark-recovery model has no parameter redundancy, but when the survival parameters are then made constant, the C/A mark-recovery model *is* then parameter redundant. In this case, making the model more complex allows parameters to be estimated, but these may be poor estimates for the underlying values of the parameters as there is less amount of information about each parameter than before. We can also see from the table that every model apart from when the detection parameters are season-, survey- and state-dependent, is full rank for the full data set including the occupancy-histories with missing observations. This is a surprising result as it is a relatively small data set with only 178 distinct occupancy-histories, some of which have missing observations too. □

Table 6.3: Parameter redundancies for multiple-states occupancy models exploring the green frogs data set of MacKenzie et al. (2009)

Model	Number of parameters	Deficiency for only full histories	Deficiency including all partial histories
$\varphi(\cdot)\ p(\cdot)$	2	0	0
$\varphi(\cdot)\ p(t)$	6	0	0
$\varphi(\cdot)\ p(j)$	4	0	0
$\varphi(\cdot)\ p(t,j)$	16	0	0
$\varphi(\cdot)\ p(m)$	7	0	0
$\varphi(\cdot)\ p(t,m)$	31	3	0
$\varphi(\cdot)\ p(j,m)$	19	1	0
$\varphi(\cdot)\ p(t,j,m)$	91	50	15
$\varphi(t)\ p(\cdot)$	6	0	0
$\varphi(t)\ p(t)$	10	0	0
$\varphi(t)\ p(j)$	8	0	0
$\varphi(t)\ p(t,j)$	20	0	0
$\varphi(t)\ p(m)$	11	0	0
$\varphi(t)\ p(t,m)$	35	7	0
$\varphi(t)\ p(j,m)$	23	1	0
$\varphi(t)\ p(t,j,m)$	95	54	19
$\varphi(m)\ p(\cdot)$	16	0	0
$\varphi(m)\ p(t)$	20	0	0
$\varphi(m)\ p(j)$	18	0	0
$\varphi(m)\ p(t,j)$	30	0	0
$\varphi(m)\ p(m)$	21	0	0
$\varphi(m)\ p(t,m)$	45	0	0
$\varphi(m)\ p(j,m)$	33	1	0
$\varphi(m)\ p(t,j,m)$	105	38	7
$\varphi(t,m)\ p(\cdot)$	52	1	0
$\varphi(t,m)\ p(t)$	56	2	0
$\varphi(t,m)\ p(j)$	54	0	0
$\varphi(t,m)\ p(t,j)$	66	3	0
$\varphi(t,m)\ p(m)$	57	1	0
$\varphi(t,m)\ p(t,m)$	81	19	0
$\varphi(t,m)\ p(j,m)$	69	3	0
$\varphi(t,m)\ p(t,j,m)$	141	74	9

6.2 Including a species original occupancy state

An extension to the multiple-state model in Section 6.1 explores the case where the species' original state can also be considered in the model. This may be of use when a study is begun at a certain point when the experimenter knows that the species is definitely present at the very start of the study. The reason why this model is advantageous to use is by considering when the occupancy probabilities are not season-dependent. If the occupancy parameters are not season-dependent, this means in this model we can define the parameters $\varphi_0^{[1]}$ and $\varphi_c^{[1,1]}$ to be the same. This then reduces the number of total parameters in the model. Under this extension, the row vector $\phi_{0,s}$ becomes

$$\phi_{0,a,s} = \left[\begin{array}{ccccccc} \varphi_{0,s}^{[a,m]} & \varphi_{0,s}^{[a,m-1]} & \cdots & \varphi_{0,s}^{[a,2]} & \varphi_{0,s}^{[a,1]} & 1 - \sum_{l=1}^{m} \varphi_{0,s}^{[a,l]} \end{array} \right], \tag{6.5}$$

where a indicates the state of the particular occupancy-history at time 0. The other notation from Section 6.1.3 remains the same. The previous state for each occupancy-history can also be distinct. This information on the initial species occupation state is given in the `Maple` procedure `occmodstatesfirsthist` in the vector **S3**. There can be mix of occupancy-histories which have original state known and unknown as well. When an occupancy-history has no known original state, the equivalent row in **S3** has the number '-1' to denote this. This means that when the information about the original state of an occupancy-history is known, it is incorporated into the model as $\varphi_{0,s}^{[a,m]}$ for known state a. The occupancy-histories where the original state is unknown changes the notation, as seen previously in this thesis, to become $\varphi_{0,s}^{[-1,m]}$, but this is equivalent to the previous notation of $\varphi_{0,s}^{[m]}$. Further note that if the transition probabilities are defined as not being state-dependent, the original state is not required in the input and the previous procedure `occmodstategroupshist` should be used from Example 6.9 in Section 6.1.4.

Example 6.11 - The multiple-states occupancy model with an initial state: [See electronic appendix `example6.11.mw`] Consider a data set containing 9 occupancy-histories where there are two surveys in the first season and three surveys in the second season with two different observational states: 11|100(U), 10|100(U), 10|010(1), 10|001(1), 01|111(U), 01|000(1), 00|011(0), 00|001(U) and 00|000(0), where (U) indicates the occupancy-history has an unknown original state and (a) where the species is originally in state a for $a = \{0, 1\}$. The parameters in the multiple-states model with group-dependency can be constant, season- and/or state-dependent and/or group-dependent,

with the detection probabilities also being possibly survey-dependent. We have the inputs **T** as the occupancy-histories, **S** as the number of surveys per season, **S2** as the column vector denoting which group the occupancy-history in **T** is from, and **S3** as the column vector denoting which original state the occupancy-history in **T** was in. As there are only two different states in this model the options for the values in **S3** are '1' if the species was originally present at that site, '0' if the species was originally absent at that site, or '-1' if the species' original state was unknown. If we do not consider any group-dependence in this example we can then code the model as follows: The inputs are given as

$$\mathbf{S} = \begin{bmatrix} 2 & 3 \end{bmatrix},$$

$$\mathbf{S2} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T,$$

$$\mathbf{S3} = \begin{bmatrix} -1 & -1 & 1 & 1 & -1 & 1 & 0 & -1 & 0 \end{bmatrix}^T,$$

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consider the two-state model where the occupancy parameters are only state-dependent and the detection parameters are season- and survey-dependent. We do not consider any state-dependent or group-dependent parameters in this example. We can then obtain the extrinsic parameter redundancy corresponding to this data set by the following `Maple` code:

```
S := < <2|3> >;
S2 := Matrix(9,1,1);
# Observe that we still need this vector even if
#  there is no group-dependence to be considered.
S3 := < <-1>, <-1>, <1>, <1>, <-1>, <1>, <0>, <-1>, <0> >;
# '-1' denotes an unknown original state.
T := < <1|1|1|0|0>, <1|0|1|0|0>, <1|0|0|1|0>, <1|0|0|0|1>,
<0|1|1|1|1>, <0|1|0|0|1>, <0|0|0|1|1>, <0|0|0|0|1>, <0|0|0|0|0> >;
```

```
kappa := occmodstatesfirsthist(2,4,7,S,S2,S3,T);
# Inputs: (M,y,z,S,S2,S3,T);
# M = number of states;
# y = transition probability; z = detection probability;
# for y: '1'=constant, '2'=season-dep, '3'=state-dep,
# '4'=season+state-dep, '5'=group-dep, '6'=season+group-dep,
# '7'=state+group-dep, '8'=season+state+group-dep;
#  for z: '1'=constant, '2'=season-dep, '3'=survey-dep,
# '4'=season+survey-dep, '5'=state-dep, '6'=season+state-dep,
# '7'=survey+state-dep, '8'=season+survey+state-dep,
# '9'=group-dep, '10'=season+group-dep, '11'=survey+group-dep,
# '12'=season+survey+group-dep, '13'=state+group-dep,
# '14'=season+state+group-dep, '15'=survey+state+group-dep,
# '16'=season+survey+state+group-dep;
# S = row vector of the number of surveys per season;
# S2 = column vector determining the group of each occupancy-history;
# S3 = column vector determining the original state of each
#  occupancy-history in T respectively; if the entry in S3 is equal
#  to -1 then this shows the state at time 0 was unknown;
# T = list of all occupancy-histories.
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

This generates the exhaustive summary, where the constant group-dependent subscripts are omitted, as

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_{c,-1,1}p_{1,1,c,c}p_{1,2,c,c}\varphi_{c,1,1}p_{2,1,c,c}\bar{p}_{2,2,c,c}\bar{p}_{2,3,c,c} \\ \varphi_{c,-1,1}p_{1,1,c,c}\bar{p}_{1,2,c,c}\varphi_{c,1,1}p_{2,1,c,c}\bar{p}_{2,2,c,c}\bar{p}_{2,3,c,c} \\ \varphi_{c,1,1}^2 p_{1,1,c,c}\bar{p}_{1,2,c,c}\bar{p}_{2,1,c,c}p_{2,2,c,c}\bar{p}_{2,3,c,c} \\ \varphi_{c,1,1}^2 p_{1,1,c,c}\bar{p}_{1,2,c,c}\bar{p}_{2,1,c,c}\bar{p}_{2,2,c,c}p_{2,3,c,c} \\ \varphi_{c,-1,1}\bar{p}_{1,1,c,c}p_{1,2,c,c}\varphi_{c,1,1}p_{2,1,c,c}p_{2,2,c,c}p_{2,3,c,c} \\ \varphi_{c,1,1}\bar{p}_{1,1,c,c}p_{1,2,c,c}\left\{\varphi_{c,1,1}\bar{p}_{2,1,c,c}\bar{p}_{2,2,c,c}\bar{p}_{2,3,c,c} + \bar{\varphi}_{c,1,1}\right\} \\ \left\{\varphi_{c,0,1}\bar{p}_{1,1,c,c}\bar{p}_{1,2,c,c}\varphi_{c,1,1} + \bar{\varphi}_{c,0,1}\varphi_{c,0,1}\right\}\bar{p}_{2,1,c,c}p_{2,2,c,c}p_{2,3,c,c} \\ \left\{\varphi_{c,-1,1}\bar{p}_{1,1,c,c}\bar{p}_{1,2,c,c}\varphi_{c,1,1} + \bar{\varphi}_{c,-1,1}\varphi_{c,0,1}\right\}\bar{p}_{2,1,c,c}\bar{p}_{2,2,c,c}p_{2,3,c,c} \\ \left\{\varphi_{c,0,1}\bar{p}_{1,1,c,c}\bar{p}_{1,2,c,c}\varphi_{c,1,1} + \bar{\varphi}_{c,0,1}\varphi_{c,0,1}\right\}\bar{p}_{2,1,c,c}\bar{p}_{2,2,c,c}\bar{p}_{2,3,c,c}\cdots \\ \cdots + \varphi_{c,0,1}\bar{p}_{1,1,c,c}\bar{p}_{1,2,c,c}\bar{\varphi}_{c,1,1} + \bar{\varphi}_{c,0,1}^2 \end{bmatrix},$$

where $\varphi_{c,l,m} = \varphi_c^{[l,m]}$ and $p_{t,j,c,c} = p_{t,j}^{c,c}$ for survey j in season t with observed state l at true state m (there is no state-dependency on the detection parameters), where $\bar{\varphi}_{c,1,1} = 1 - \varphi_{c,1,1}$. The parameter set of this model is $\boldsymbol{\theta} = [p_{1,1,c,c}, p_{1,2,c,c}, p_{2,1,c,c}, p_{2,2,c,c}, p_{2,3,c,c}, \varphi_{c,-1,1}, \varphi_{c,0,1}, \varphi_{c,1,1}]^T$. The derivative matrix is not given here to save space but can be found in the `Maple` file `example6.11.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$$\mathbf{r} := \mathbf{9}, \mathbf{d} := \mathbf{0}.$$

This means that including information about original states at time 0 does not cause parameter redundancy in this case. We can now estimate more parameters in this 2 state model of the form $\varphi_{0,s}^{[a,m]}$ for $a = \{-1, 0, 1\}$ with -1 indicating an unknown original state, rather than just the sole $\varphi_{0,s}^{[m]}$ parameter. Therefore, if this information regarding original state is available, then it can be beneficial to use this information in this procedure to examine parameter redundancies. There can also be multiple occupancy-histories with different original states, e.g. three different exhaustive summary terms can be made from the observations $00|00(U)$, $00|00(1)$ and $00|00(0)$ due to different original states. This extension can then help produce better point estimates for the $\varphi_{0,s}^{[l,m]}$ parameters in the model. □

6.3 The two-species interaction occupancy model

We have explored a variety of occupancy models with a view to estimating the probability of occupancy in various studies. However, a limitation in these models in that you can only examine a single species. Multiple species are not accounted for in the same model. It is possible to consider each species separately, but this would not be ideal if some assumptions lead you to believe that the presence or absence of one species has a direct interaction on the presence or absence of another species. So therefore this **two-species interaction occupancy model** explores two different interacting species being considered at the same time. We examine extrinsic parameter redundancy in the occupancy model from MacKenzie et al. (2004) where they consider a single-season-only model, however we use the extension of this model to account for multiple seasons given in MacKenzie et al. (2006, Chapter 8). We also illustrate how a model with multiple interacting species could be formed later on in Section 6.3.2. The work from MacKenzie et al. (2004, 2006) has been used to analyse a variety of

interacting animal species, including northern spotted and barred owls in Bailey et al. (2009), Virginia rail and California Black rail waterbirds in Richmond et al. (2010), and interacting snake species in Steen et al. (2014).

6.3.1 Model formulation

The two-species interaction occupancy model has different sets of parameters to account for different species, where we denote the two species as 'species A' and 'species B'. The important consideration to make in this model is the fact we allow the presence/absence of species A to have an effect on the presence/absence of species B. If there is no interaction between two species it would be possible to consider a simpler occupancy model such as from Section 5.7, but we assume here that species A does has an effect on species B. The model parameters are as follows:

- $\varphi_{0,s}^{[AB]}$: The probability of both species being present at time $t = 0$.
- $\varphi_{0,s}^{[A]}$: The probability that species A is present regardless of the status of species B at time $t = 0$.
- $\varphi_{0,s}^{[B]}$: The probability that species B is present regardless of the status of species A at time $t = 0$.
- $\varphi_{t,s}^{[X,Y]}$: The transition probability of going from occupancy category X in season $t-1$ to occupancy category Y in season t, for any season $t = 1, \ldots, T-1$. X and Y can be either 'AB' where both species are present, 'A' where only species A is present, 'B' where only species B is present, or 'U' where both species are absent.
- $r_{t,j,s}^{AB}$: The probability of detecting both species, given both species are present at survey j.
- $r_{t,j,s}^{Ab}$: The probability of detecting species A but not species B, given both species are present.
- $r_{t,j,s}^{aB}$: The probability of detecting species B but not species A, given both species are present.
- $p_{t,j,s}^{A}$: The probability of detecting species A, given only species A is present.
- $p_{t,j,s}^{B}$: The probability of detecting species B, given only species B is present.

Every parameter in this model can be season-dependent and/or group-dependent, with the detection probabilities also possibly being survey-dependent. Season-dependence

is denoted by the subscript t, group-dependence with the subscript s, and survey-dependence with the subscript j as is commonly used in this chapter. Observe that the transition probability of going from state X to the unknown state U is equal to $\varphi_{t,s}^{[X,U]} = 1 - \varphi_{t,s}^{[X,AB]} - \varphi_{t,s}^{[X,A]} - \varphi_{t,s}^{[X,B]}$, and the probability that detecting neither species given both species are present is equal to $r_{t,j,s}^{ab} = 1 - r_{t,j,s}^{AB} - r_{t,j,s}^{Ab} - r_{t,j,s}^{aB}$, thus reducing the number of parameters in the model. We also change our parameter notation compared to that given in MacKenzie et al. (2006, pp.245) as they use the parameters $\{\epsilon, \nu, \gamma, \eta, \omega\}$ as an alternative to our notation of $\varphi_{t,s}^{[X,Y]}$, which gives the transition probability between various states of extinction and colonisation for all different possibilities of species presence/absence. There is also a different parameterisation of this two-species model given in Richmond et al. (2010). These alternative parameterisations will have the same extrinsic parameter redundancy results as can be shown using the reparameterisation theorem of Theorem 2.3.

There are some parameters that are of particular interest to experimenters performing two-species occupancy studies. These include the transition probabilities $\varphi_{t,s}^{[A,AB]}$ and $\varphi_{t,s}^{[B,AB]}$, which are the probabilities that both species are present at the study location in season t given one of the species was absent in season $t-1$. The parameters $\varphi_{t,s}^{[A,B]}$ and $\varphi_{t,s}^{[B,A]}$ are also of interest where the two species switch occupancy states as these parameters examine the relationships between native and invasive species. Commonly in work on single-season occupancy studies, experimenters instead wish to obtain levels of co-occurrence between the two interacting species in the study. This co-occurrence ratio is called the species interaction factor, which is denoted as ϕ, but is not to be confused between the survival parameters from previous chapters. This species interaction factor is given as

$$\phi = \frac{\varphi^{[AB]}}{\varphi^{[A]}\varphi^{[B]}}. \tag{6.6}$$

If the two species occupy the study location independently, then $\varphi^{[AB]} = \varphi^{[A]} \times \varphi^{[B]}$ and therefore $\phi = 1$. A species interaction factor value of less than one suggests they co-occur at the location less regularly then they would do if the species are independent, and a species interaction factor value of greater than one suggests they co-occur at the location more regularly.

We now write down a general form for any occupancy-history in this two-species interaction occupancy model. Let $\phi_{0,s}$ be the initial transition row vector for group s

given as

$$\phi_{0,s} = \left[\begin{array}{cccc} \varphi_{0,s}^{AB} & \varphi_{0,s}^{A} - \varphi_{0,s}^{AB} & \varphi_{0,s}^{B} - \varphi_{0,s}^{AB} & 1 - \varphi_{0,s}^{A} - \varphi_{0,s}^{B} + \varphi_{0,s}^{AB} \end{array} \right]. \tag{6.7}$$

The later transition matrices are denoted as $\phi_{t,s}$ for seasons $t = 1, \ldots, T-1$ for group s, and are given by

$$\phi_{t,s} = \left[\begin{array}{cccc} \varphi_{t,s}^{[AB,AB]} & \varphi_{t,s}^{[AB,A]} & \varphi_{t,s}^{[AB,B]} & 1 - \varphi_{t,s}^{[AB,AB]} - \varphi_{t,s}^{[AB,A]} - \varphi_{t,s}^{[AB,B]} \\ \varphi_{t,s}^{[A,AB]} & \varphi_{t,s}^{[A,A]} & \varphi_{t,s}^{[A,B]} & 1 - \varphi_{t,s}^{[A,AB]} - \varphi_{t,s}^{[A,A]} - \varphi_{t,s}^{[A,B]} \\ \varphi_{t,s}^{[B,AB]} & \varphi_{t,s}^{[B,A]} & \varphi_{t,s}^{[B,B]} & 1 - \varphi_{t,s}^{[B,AB]} - \varphi_{t,s}^{[B,A]} - \varphi_{t,s}^{[B,B]} \\ \varphi_{t,s}^{[U,AB]} & \varphi_{t,s}^{[U,A]} & \varphi_{t,s}^{[U,B]} & 1 - \varphi_{t,s}^{[U,AB]} - \varphi_{t,s}^{[U,A]} - \varphi_{t,s}^{[AB,B]} \end{array} \right]. \tag{6.8}$$

The column vector of the detection probabilities is denoted by $\mathbf{p}_{t,s}^{\{h_{X,t,s}^{A}\},\{h_{X,t,s}^{B}\}}$ for the Xth occupancy-history $h_{X,t,s}^{A}$ of species A and for the Xth occupancy-history $h_{X,t,s}^{B}$ of species B in season t for group s. The column vector $\mathbf{p}_{t,s}^{\{h_{X,t,s}^{A}\},\{h_{X,t,s}^{B}\}}$ is a 4-entry column vector with the probability of both species being present in the first row, the probability of only species A being present in the second row, the probability of only species B being present in the third row, and the probability of both species being absent in the final row. This means that if both species are detected then the vector would only have a non-zero entry in the first row. If only species A was detected there would only be non-zero entries in the first two rows of the vector, and if only species B was detected there would only be non-zero entries in the first and third rows of the vector. We only consider species which have the two states as presence and absence for these two-species interaction models. If we let

$$\delta_{t,j,s}^{A} = \begin{cases} 1 & \text{when species A is detected on survey } j \text{ in season } j \text{ for group } s, \\ 0 & \text{when species A is not detected on survey } j \text{ in season } j \text{ for group } s, \end{cases}$$

for species A and

$$\delta_{t,j,s}^{B} = \begin{cases} 1 & \text{when species B is detected on survey } j \text{ in season } j \text{ for group } s, \\ 0 & \text{when species B is not detected on survey } j \text{ in season } j \text{ for group } s, \end{cases}$$

for species B during season t at survey j for group s, we can define this vector $\mathbf{p}_{X,t,s}^{\{h_{X,t,s}^A\},\{h_{X,t,s}^B\}}$, to be given as

$$\left[\begin{array}{c} \prod_{j=1}^{k_t}\left\{\delta_{t,j,s}^A\delta_{t,j,s}^B r_{t,j,s}^{AB}+\delta_{t,j,s}^A\bar{\delta}_{t,j,s}^B r_{t,j,s}^{Ab}+\bar{\delta}_{t,j,s}^A\delta_{t,j,s}^B r_{t,j,s}^{aB}+\bar{\delta}_{t,j,s}^A\bar{\delta}_{t,j,s}^B\left(1-r_{t,j,s}^{AB}-r_{t,j,s}^{Ab}-r_{t,j,s}^{aB}\right)\right\} \\ \left\{\prod_{j=1}^{k_t}\bar{\delta}_{t,j,s}^B\right\}\left(\delta_{t,j,s}^A p_{t,j,s}^A+\bar{\delta}_{t,j,s}^A\bar{p}_{t,j,s}^A\right) \\ \left\{\prod_{j=1}^{k_t}\bar{\delta}_{t,j,s}^A\right\}\left(\delta_{t,j,s}^B p_{t,j,s}^B+\bar{\delta}_{t,j,s}^B\bar{p}_{t,j,s}^B\right) \\ \prod_{j=1}^{k_t}\bar{\delta}_{t,j,s}^A\bar{\delta}_{t,j,s}^B \end{array}\right], \tag{6.9}$$

where $\bar{\delta}_{t,j,s}^A=(1-\delta_{t,j,s}^A)$. We show how occupancy-histories for the two-species interaction model can be represented in Examples 6.12 to 6.14 below.

Example 6.12: Consider an example where there are three surveys in season t, where the occupancy-history for species A is observed as $h^A=110$ and the occupancy-history for species B is observed as $h^B=100$. The detection probability vector is then given as

$$\mathbf{p}_{X,t,s}^{\{h_{110,t,s}^A\},\{h_{100,t,s}^B\}}=\left[\begin{array}{c} r_{t,1,s}^{AB}r_{t,2,s}^{Ab}(1-r_{t,3,s}^{AB}-r_{t,3,s}^{Ab}-r_{t,3,s}^{aB}) \\ 0 \\ 0 \\ 0 \end{array}\right].$$

□

Example 6.13: Consider the occupancy-history where the occupancy-history for species A is observed as $h^A=000$ and the occupancy-history for species B is observed as $h^B=011$. This gives the following column vector for the detections

$$\mathbf{p}_{X,t,s}^{\{h_{000,t,s}^A\},\{h_{011,t,s}^B\}}=\left[\begin{array}{c} (1-r_{t,1,s}^{AB}-r_{t,1,s}^{Ab}-r_{t,1,s}^{aB})r_{t,2,s}^{aB}r_{t,3,s}^{aB} \\ 0 \\ (1-p_{t,1,s}^B)p_{t,2,s}^B p_{t,3,s}^B \\ 0 \end{array}\right].$$

□

Example 6.14: Consider the occupancy-history where the occupancy-history for species A is observed as $h^A = 000$ and the occupancy-history for species B is observed as $h^B = 000$, which results in the column vector

$$\mathbf{p}_{X,t,s}^{\{h^A_{000,t,s}\},\{h^B_{000,t,s}\}} = \begin{bmatrix} \prod_{j=1}^{3}(1 - r^{AB}_{t,j,s} - r^{Ab}_{t,j,s} - r^{aB}_{t,j,s}) \\ \prod_{j=1}^{3}(1 - p^{A}_{t,j,s}) \\ \prod_{j=1}^{3}(1 - p^{B}_{t,j,s}) \\ 1 \end{bmatrix}.$$

□

We denote $\Delta(\mathbf{p}_{X,t,s})$ as the matrix where we diagonalise the $\mathbf{p}_{X,t,s}$ vector by putting the elements of $\mathbf{p}_{X,t,s}$ on the diagonal and zeros on the non-diagonal elements of the matrix. The probability of a certain occupancy-history occurring is expressed as

$$Pr(h_X) = \phi_{0,s}\left[\prod_{t=1}^{T-1}\Delta\left(\mathbf{p}_{X,t,s}^{\{h^A_{X,t,s}\},\{h^B_{X,t,s}\}}\right)\phi_{t,s}\right]\mathbf{p}_{X,T,s}^{\{h^A_{X,T,s}\},\{h^B_{X,T,s}\}}. \tag{6.10}$$

This gives the likelihood

$$L = \prod_{i=1}^{N} Pr(h_X),$$

for the N different occupancy-histories observed in a study. We give Example 6.15 below to show how the probability of a single occupancy-history with multiple seasons can obtained.

Example 6.15: Consider the occupancy-history where the occupancy-history for species A is observed as $h^A = 10|00$ and the occupancy-history for species B is observed as $h^B = 01|00$. The probability of this occupancy-history with no group-dependency can

be obtained by

$$
\begin{aligned}
Pr(h) &= \phi_0 \Delta \left(\mathbf{p}_{X,1,c}^{\left\{h_{10,1,c}^{A}\right\},\left\{h_{01,1,c}^{B}\right\}} \right) \phi_1 \mathbf{p}_{X,2,c}^{\left\{h_{00,2,c}^{A}\right\},\left\{h_{00,2,c}^{B}\right\}} \\
&= \begin{bmatrix} \varphi_{0,c}^{[AB]} \\ \varphi_{0,c}^{[A]} - \varphi_{0,c}^{[AB]} \\ \varphi_{0,c}^{[B]} - \varphi_{0,c}^{[AB]} \\ 1 - \varphi_{0,c}^{[A]} - \varphi_{0,c}^{[B]} + \varphi_{0,c}^{[AB]} \end{bmatrix}^T \cdot \begin{bmatrix} r_{1,1,c}^{Ab} r_{1,2,c}^{aB} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} \varphi_{1,c}^{[AB,AB]} & \varphi_{1,c}^{[AB,A]} & \varphi_{1,c}^{[AB,B]} & 1 - \varphi_{1,c}^{[AB,AB]} - \varphi_{1,c}^{[AB,A]} - \varphi_{1,c}^{[AB,B]} \\ \varphi_{1,c}^{[A,AB]} & \varphi_{1,c}^{[A,A]} & \varphi_{1,c}^{[A,B]} & 1 - \varphi_{1,c}^{[A,AB]} - \varphi_{1,c}^{[A,A]} - \varphi_{1,c}^{[A,B]} \\ \varphi_{1,c}^{[B,AB]} & \varphi_{1,c}^{[B,A]} & \varphi_{1,c}^{[B,B]} & 1 - \varphi_{1,c}^{[B,AB]} - \varphi_{1,c}^{[B,A]} - \varphi_{1,c}^{[B,B]} \\ \varphi_{1,c}^{[U,AB]} & \varphi_{1,c}^{[U,A]} & \varphi_{1,c}^{[U,B]} & 1 - \varphi_{1,c}^{[U,AB]} - \varphi_{1,c}^{[U,A]} - \varphi_{1,c}^{[AB,B]} \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} (1 - r_{2,1,c}^{AB} - r_{2,1,c}^{Ab} - r_{2,1,c}^{aB})(1 - r_{2,2,c}^{AB} - r_{2,2,c}^{Ab} - r_{2,2,c}^{aB}) \\ (1 - p_{2,1,c}^{A})(1 - p_{2,2,c}^{A}) \\ (1 - p_{2,1,c}^{B})(1 - p_{2,2,c}^{B}) \\ 1 \end{bmatrix} \\
&= \varphi_{0,c}^{[AB]} r_{1,1,c}^{Ab} r_{1,2,c}^{aB} \left\{ \varphi_{1,c}^{[AB,AB]} \left(1 - r_{2,1,c}^{AB} - r_{2,1,c}^{Ab} - r_{2,1,c}^{aB}\right) \left(1 - r_{2,2,c}^{AB} - r_{2,2,c}^{Ab} - r_{2,2,c}^{aB}\right) \right. \\
&\quad + \varphi_{1,c}^{[AB,A]} \left(1 - p_{2,1,c}^{A}\right) \left(1 - p_{2,2,c}^{A}\right) + \varphi_{1,c}^{[AB,B]} \left(1 - p_{2,1,c}^{B}\right) \left(1 - p_{2,2,c}^{B}\right) \\
&\quad \left. + \left(1 - \varphi_{1,c}^{[AB,AB]} - \varphi_{1,c}^{[AB,A]} - \varphi_{1,c}^{[AB,B]}\right) \right\}.
\end{aligned}
$$

□

6.3.2 Having more than two species in the model

It is possible to extend this model beyond just two interacting species. The model however gets complex for even a small number of multiple species. To see how this model can be extended, take a three-species model for example. Define $\varphi_{0,s}^{[X]}$ as the initial occupancy state of all three of the species for group s, e.g. $\varphi_{0,s}^{[ABC]}$ indicates the probability all three species are initially present, while $\varphi_{0,s}^{[B]}$ indicates the probability

that only species B is initially present. The first transition vector can be defined as

$$\phi_{0,s} = \begin{bmatrix} \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[AB]} - \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[AC]} - \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[BC]} - \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[A]} - \varphi_{0,s}^{[AB]} - \varphi_{0,s}^{[AC]} + \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[B]} - \varphi_{0,s}^{[AB]} - \varphi_{0,s}^{[BC]} + \varphi_{0,s}^{[ABC]} \\ \varphi_{0,s}^{[C]} - \varphi_{0,s}^{[AC]} - \varphi_{0,s}^{[BC]} + \varphi_{0,s}^{[ABC]} \\ 1 - \varphi_{0,s}^{[A]} - \varphi_{0,s}^{[B]} - \varphi_{0,s}^{[C]} + \varphi_{0,s}^{[AB]} + \varphi_{0,s}^{[AC]} + \varphi_{0,s}^{[BC]} - \varphi_{0,s}^{[ABC]} \end{bmatrix}^T .$$

The transition probabilities $\varphi_{t,s}^{[X,Y]}$ for season t for group s go from occupancy category X to occupancy category Y, e.g. $\varphi_{t,s}^{[ABC,A]}$ means from going from all three species being present in season $t-1$ to season t where only species A is present. The subsequent transition matrices $\phi_{t,s}$ are then given in the matrix below

$$\begin{bmatrix} \varphi_{t,s}^{[ABC,ABC]} & \varphi_{t,s}^{[ABC,AB]} & \varphi_{t,s}^{[ABC,AC]} & \varphi_{t,s}^{[ABC,BC]} & \varphi_{t,s}^{[ABC,A]} & \varphi_{t,s}^{[ABC,B]} & \varphi_{t,s}^{[ABC,C]} & 1 - \sum_Y \varphi_{t,s}^{[ABC,Y]} \\ \varphi_{t,s}^{[AB,ABC]} & \varphi_{t,s}^{[AB,AB]} & \varphi_{t,s}^{[AB,AC]} & \varphi_{t,s}^{[AB,BC]} & \varphi_{t,s}^{[AB,A]} & \varphi_{t,s}^{[AB,B]} & \varphi_{t,s}^{[AB,C]} & 1 - \sum_Y \varphi_{t,s}^{[AB,Y]} \\ \varphi_{t,s}^{[AC,ABC]} & \varphi_{t,s}^{[AC,AB]} & \varphi_{t,s}^{[AC,AC]} & \varphi_{t,s}^{[AC,BC]} & \varphi_{t,s}^{[AC,A]} & \varphi_{t,s}^{[AC,B]} & \varphi_{t,s}^{[AC,C]} & 1 - \sum_Y \varphi_{t,s}^{[AC,Y]} \\ \varphi_{t,s}^{[BC,ABC]} & \varphi_{t,s}^{[BC,AB]} & \varphi_{t,s}^{[BC,AC]} & \varphi_{t,s}^{[BC,BC]} & \varphi_{t,s}^{[BC,A]} & \varphi_{t,s}^{[BC,B]} & \varphi_{t,s}^{[BC,C]} & 1 - \sum_Y \varphi_{t,s}^{[BC,Y]} \\ \varphi_{t,s}^{[A,ABC]} & \varphi_{t,s}^{[A,AB]} & \varphi_{t,s}^{[A,AC]} & \varphi_{t,s}^{[A,BC]} & \varphi_{t,s}^{[A,A]} & \varphi_{t,s}^{[A,B]} & \varphi_{t,s}^{[A,C]} & 1 - \sum_Y \varphi_{t,s}^{[A,Y]} \\ \varphi_{t,s}^{[B,ABC]} & \varphi_{t,s}^{[B,AB]} & \varphi_{t,s}^{[B,AC]} & \varphi_{t,s}^{[B,BC]} & \varphi_{t,s}^{[B,A]} & \varphi_{t,s}^{[B,B]} & \varphi_{t,s}^{[B,C]} & 1 - \sum_Y \varphi_{t,s}^{[B,Y]} \\ \varphi_{t,s}^{[C,ABC]} & \varphi_{t,s}^{[C,AB]} & \varphi_{t,s}^{[C,AC]} & \varphi_{t,s}^{[C,BC]} & \varphi_{t,s}^{[C,A]} & \varphi_{t,s}^{[C,B]} & \varphi_{t,s}^{[C,C]} & 1 - \sum_Y \varphi_{t,s}^{[C,Y]} \\ \varphi_{t,s}^{[U,ABC]} & \varphi_{t,s}^{[U,AB]} & \varphi_{t,s}^{[U,AC]} & \varphi_{t,s}^{[U,BC]} & \varphi_{t,s}^{[U,A]} & \varphi_{t,s}^{[U,B]} & \varphi_{t,s}^{[U,C]} & 1 - \sum_Y \varphi_{t,s}^{[U,Y]} \end{bmatrix},$$

with $\sum_Y \varphi_{t,s}^{[X,Y]}$ denoting the sum of all the transition probabilities from occupancy category X to category $Y = \{ABC, AB, AC, BC, A, B, C\}$. The detection column vector can then be defined as

$$\mathbf{p}_{X,t,s}^{\{h_{X,t,s}^A\},\{h_{X,t,s}^B\},\{h_{X,t,s}^C\}} = \begin{bmatrix} Pr(\text{All three species are present}) \\ Pr(\text{Only species A and B are present}) \\ Pr(\text{Only species A and C are present}) \\ Pr(\text{Only species B and C are present}) \\ Pr(\text{Only species A is present}) \\ Pr(\text{Only species B is present}) \\ Pr(\text{Only species C is present}) \\ Pr(\text{All species are absent}) \end{bmatrix}^T ,$$

containing detection probabilities $p^X_{t,j,s}$ with occupancy category X. The final probability of a certain occupancy-history is expressed as

$$Pr(h_X) = \phi_{0,s}\left[\prod_{t=1}^{T-1}\Delta\left(\mathbf{p}^{\{h^A_{X,t,s}\},\{h^B_{X,t,s}\},\{h^C_{X,t,s}\}}_{X,t,s}\right)\phi_{t,s}\right]\mathbf{p}^{\{h^A_{X,T,s}\},\{h^B_{X,T,s}\},\{h^C_{X,T,s}\}}_{X,T,s}. \tag{6.11}$$

As we can see though from the matrix $\phi_{t,s}$, there is a very large number of parameters in this model, so that the data source for fitting such a model needs to be substantial. Due to the large number of parameters and the difficulty of coding this model in `Maple`, even for a three-species model, we will not consider any further the occupancy model for more than two species in any depth. It is mentioned here for future reference if this model, or indeed a model with even more than three interacting species, would be applicable in another context.

6.3.3 Parameter redundancy in the two-species model

The `Maple` procedure below generates the exhaustive summary given a set of occupancy-histories using the two-species interaction occupancy model.

Example 6.16 - The two-species interaction occupancy model: [See electronic appendix `example6.16.mw`] Consider a study for a model with two seasons with two surveys per season and no group-dependency. We observe these occupancy-histories for species A: 11|11, 11|10, 10|10, 10|10, 10|10, 10|00, 01|11, 01|10, 01|10, 01|01, 01|00, 01|00, 01|00, 00|11, 00|10, 00|10, 00|10, 00|01, 00|01, 00|01, 00|01, 00|00, 00|00, 00|00, 00|00 and 00|00. We also observe these occupancy-histories for species B: 01|10, 10|11, 00|00, 00|01, 11|01, 10|00, 00|00, 00|10, 00|00, 10|10, 10|10, 00|10, 11|11, 01|10, 01|10, 00|11, 00|01, 11|10, 11|00, 01|10, 00|11, 00|10, 00|01, 10|01, 10|10 and 00|00. These occupancy-histories are given in the two matrices **TA** and **TB**, where **TA** represents the occupancy-histories of species A and **TB** represents the occupancy-histories of species B. Note that each row for occupancy-history of species A in matrix **TA** must be in the same row as the occupancy-history of species B in matrix **TB** so that it matches up between the matrices.

In the `Maple` procedure `occmodspecieshist` we need to specify **S** as the number of surveys per season and **S2** as the column vector denoting which group the occupancy-histories listed in **TA** and **TB** are from. Along with the matrices **TA** and **TB**, these inputs are given as

$$\mathbf{S} = \begin{bmatrix} 2 & 2 \end{bmatrix},$$

$$\mathbf{S2} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix}^T,$$

$$\mathbf{TA} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

$$\mathbf{TB} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Consider a model with season-dependent occupancy parameters and season- and survey-dependent detection parameters. We can then obtain the extrinsic parameter redundancy results corresponding this data set by the following `Maple` code:

```
S := < <2|2> >;
S2 := Matrix(26,1,1);
# a 1 by 26 column vector with all entries as 1: Observe that
#  this vector still needs to be defined even if there is no
#  parameters which are group-dependent.
TA := < <1|1|1|1>, <1|1|1|0>, <1|0|1|0>, <1|0|1|0>, ..., <0|0|0|0> >;
TB := < <0|1|1|0>, <1|0|1|1>, <0|0|0|0>, <0|0|0|1>, ..., <0|0|0|0> >;
kappa := occmodspecieshist(2,4,S,S2,TA,TB);
# Inputs: (y,z,S,S2,TA,TB);
# y = transition probability; z = detection probability;
# for y: '1'=constant, '2'=season-dep, '3'=state-dep,
# '4'=season+state-dep; for z: '1'=constant, '2'=season-dep,
# '3'=survey-dep, '4'=season+survey-dep, '5'=state-dep,
# '6'=season+state-dep, '7'=survey+state-dep,
# '8'=season+survey+state-dep;
# S = row vector of the number of surveys per season;
# S2 = column vector to determine the group of each occupancy-history;
# TA = list of all occupancy-histories for species A;
```

```
# TB = list of all occupancy-histories for species B.
theta := parsproc(kappa):
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

This generates the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} \varphi_{0,c,AB} r_{1,1,c,Ab} r_{1,2,c,AB} \varphi_{1,c,AB,AB} r_{2,1,c,AB} r_{2,2,c,Ab} \\ \varphi_{0,c,AB} r_{1,1,c,AB} r_{1,2,c,Ab} \varphi_{1,c,AB,AB} r_{2,1,c,AB} r_{2,2,c,aB} \\ \varphi_{0,c,AB} r_{1,1,c,AB} r_{1,2,c,aB} \left\{ \varphi_{1,c,AB,AB} r_{2,1,c,Ab} r_{2,2,c,Ab} + \varphi_{1,c,AB,A} p_{2,1,c,A} p_{2,2,c,A} \right\} \\ \left\{ \varphi_{0,c,AB} r_{1,1,c,Ab} (1 - r_{1,2,c,AB} - r_{1,2,c,Ab} - r_{1,2,c,aB}) \varphi_{1,c,AB,AB} \ \cdots \right. \\ \left. + (\varphi_{0,c,A} - \varphi_{0,c,AB}) p_{1,1,c,A} (1 - p_{1,2,c,A}) \varphi_{1,c,A,AB} \right\} r_{2,1,c,Ab} r_{2,2,c,aB} \\ \varphi_{0,c,AB} r_{1,1,c,AB} r_{1,2,c,aB} \varphi_{1,c,AB,AB} r_{2,1,c,Ab} r_{2,2,c,aB} \\ \vdots \end{bmatrix},$$

where the transition probabilities are $\varphi_{t,c,X,Y} = \varphi_{t,c}^{[X,Y]}$ and $\varphi_{0,c,X} = \varphi_{0,c}^{[Y]}$ for $X = \{U, A, B, AB\}$ and $Y = \{A, B, AB\}$, and the detection probabilities are $r_{t,j,c,V} = r_{t,j,c}^{V}$ and $p_{t,j,c,W} = p_{t,j,c}^{W}$ for $V = \{AB, Ab, aB\}$ and $W = \{A, B\}$. The parameter set and derivative matrix are not given here to save space but can be found in the `Maple` file `example6.16.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{26}, \mathbf{d} := \mathbf{9}$.

This has a parameter deficiency of 9 with 26 estimable parameter combinations. However, if we fixed the detection probabilities to be constant for example, then there is an extrinsic parameter deficiency of only 2 with 18 estimable parameter combinations.

□

6.3.4 A real data example: Different species in the Yellowstone and Grand Teton data set

We now illustrate extrinsic parameter redundancy results for two-species interaction occupancy models using the `Maple` procedure `occmodspecieshist` from Example 6.16 on the amphibian breeding in Yellowstone and Grand Teton national parks data set of Gould et al. (2012).

Example 6.17 - The two-species interaction occupancy model for the Yellowstone and Grand Teton data set: To show a real data set using the two-species interaction model, we examine the Yellowstone and Grand Teton data set from Gould et al. (2012) which we previously explored in Sections 5.7 and 5.8. The data considers occupancy-histories from three different species on tiger salamanders (*Ambystoma tigrinum*), Columbian spotted frogs (*Rana luteiventris*) and boreal chorus frogs (*Pseudacris maculata*). We could use the model from Section 6.3.2 for a three-species interaction model, but will only explore two-species interaction models in this thesis. To further reinforce that the code accommodates group-dependence as well, we propose to include group-dependence by splitting the data into four different groups: Group 1 examines the observations from the wetland sites, group 2 examines the observations obtained from the first five site ID numbers in the study, group 3 examines the observations obtained from the sixth to the tenth site ID numbers in the study, and group 4 comprises the remaining observations in the study. This is again an arbitrarily chosen set of groups not picked by any sort of contextual background, but this does mean groups 2, 3 and 4 roughly have the same number of occupancy-histories. We also consider the occupancy-histories with missing observations for this analysis. The parameters in the model as before can be season-dependent, (t), and/or group-dependent, (s), as well as possibly being survey-dependent for the detection probabilities, (j). We also examine three different parameter deficiencies for each of the species comparisons we can make, i.e. we compare tiger salamanders to Columbian spotted frogs, Columbian spotted frogs to boreal chorus frogs, and boreal chorus frogs to tiger salamanders. This analysis is presented in Table 6.4.

We observe from Table 6.4 that most of the simpler two-species interaction models remain full rank for this data set. The data does however have large parameter deficiencies when the occupancy parameters are both season- and group-dependent. This is due to the number of parameters when these two dependencies are factored in, as there is a total of $\{12(T-1)+3\}S$ different occupancy parameters when they are season- and group-dependent (which is equal to 156 different parameters in this study). There are also some differences when you compare the species pairs, such as there are lower parameter deficiencies when comparing the Columbian spotted frogs and boreal chorus frogs then comparing either of them with the tiger salamanders: This is the effect of slightly different confounding parameters between the different cases, possibly due to large number of missing observations for the tiger salamanders. □

Table 6.4: Parameter redundancies in two-species interaction occupancy models exploring the Yellowstone and Grand Teton data set of Gould et al. (2012)

Model	Number of parameters	Parameter deficiency of Salamanders & Columbian frogs	Parameter deficiency of Columbian frogs & chorus frogs	Parameter deficiency of Chorus frogs & salamanders
$\varphi(\cdot)\ p(\cdot)$	20	0	0	0
$\varphi(\cdot)\ p(t)$	35	0	0	0
$\varphi(\cdot)\ p(j)$	25	0	0	0
$\varphi(\cdot)\ p(t,j)$	55	0	0	0
$\varphi(\cdot)\ p(s)$	35	0	0	0
$\varphi(\cdot)\ p(t,s)$	95	1	0	1
$\varphi(\cdot)\ p(j,s)$	55	0	0	0
$\varphi(\cdot)\ p(t,j,s)$	175	28	13	32
$\varphi(t)\ p(\cdot)$	44	0	0	0
$\varphi(t)\ p(t)$	59	0	0	0
$\varphi(t)\ p(j)$	49	0	0	0
$\varphi(t)\ p(t,j)$	79	0	0	0
$\varphi(t)\ p(s)$	59	0	0	0
$\varphi(t)\ p(t,s)$	118	1	0	1
$\varphi(t)\ p(j,s)$	79	0	0	0
$\varphi(t)\ p(t,j,s)$	199	29	13	47
$\varphi(s)\ p(\cdot)$	65	0	0	0
$\varphi(s)\ p(t)$	80	0	0	0
$\varphi(s)\ p(j)$	70	0	0	0
$\varphi(s)\ p(t,j)$	100	0	0	0
$\varphi(s)\ p(s)$	80	0	0	0
$\varphi(s)\ p(t,s)$	140	7	9	14
$\varphi(s)\ p(j,s)$	100	0	0	1
$\varphi(s)\ p(t,j,s)$	220	53	40	74
$\varphi(t,s)\ p(\cdot)$	161	20	17	31
$\varphi(t,s)\ p(t)$	176	20	17	31
$\varphi(t,s)\ p(j)$	166	20	17	31
$\varphi(t,s)\ p(t,j)$	196	21	17	44
$\varphi(t,s)\ p(s)$	176	24	19	42
$\varphi(t,s)\ p(t,s)$	236	67	54	86
$\varphi(t,s)\ p(j,s)$	196	34	25	56
$\varphi(t,s)\ p(t,j,s)$	316	134	112	164
A three-species interaction occupancy model is not considered here.				

6.4 The Royle-Nichols model

6.4.1 Obtaining abundance estimates from occupancy models

The final work in this chapter explores the **Royle-Nichols occupancy model** proposed by Royle and Nichols (2003), which explores the analysis of animal abundance in the species population. In previous occupancy models we do not allow abundance to influence detectability. In some examples this may be a valid assumption, but there are many examples of where a higher species abundance means there is a higher species detection probability. Royle and Nichols (2003) exploit the link between the variation in abundance and the variation in detection probability to propose a model which considers heterogeneity in detection probability as a way of viewing heterogeneity in a species' abundance. The Royle-Nichols model allows the estimation of species abundance from repeated surveys of species detection with the advantage of us not needing to uniquely mark individuals. Examples of the use of the Royle-Nichols models include examining small mammals in tracking tube data from Nams and Gillis (2003), examining green frogs in Royle (2004a), and examining mountain plovers in Dreitz et al. (2006).

Royle and Nichols (2003) use the model of Section 5.3 and MacKenzie et al. (2002) as a basis for their model. The Royle-Nichols model only contains a single 'season' with multiple surveys in the study. We let n be the number of species individuals that are present at the study location. Further let r be the *individual* detection probability. Note that r is the individual detection probability rather than the species detection probability given as p in our previous occupancy models. These are two different parameters which have different meanings as there can be multiple individuals in a single species which could be detected. The probability of detecting a species in a survey given the species is present is

$$p = 1 - (1 - r)^n. \tag{6.12}$$

This probability is calculated as one minus the probability of not detecting that species, where the probability of not detecting that species is equal to the probability of not detecting any of the individuals in the species, $(1 - r)^n$. We note that p may be equal to zero, which is the case when the probability of individual detection is $r = 0$ or when there are no individuals as $n = 0$. It is also seen that if $n = 1$, then the probability of species detection is exactly equal to the probability of individual detection, $p = r$.

If we let w be the number of detections observed in k repeated surveys, then a likelihood for the Royle-Nichols model, conditional on the number of n individuals, can be expressed as

$$L(w|n) = \binom{k}{w} \{1-(1-r)^n\}^w \{(1-r)^n\}^{k-w}, \tag{6.13}$$

where $\binom{k}{w}$ is the binomial coefficient of there being w detections from k surveys. If we further consider a likelihood which is unconditional on the number of n individuals at the study location, this can be given as

$$L(w) = \sum_{n=0}^{\infty} \binom{k}{w} \{1-(1-r)^n\}^w \{(1-r)^n\}^{k-w} \cdot f(n), \tag{6.14}$$

with a mixing distribution $f(n)$. This mixing distribution gives the probability that n individuals are present at the study location. A commonly used distribution for $f(n)$ is the Poisson distribution with

$$f(n) = \frac{e^{-\mu}\mu^n}{n!}. \tag{6.15}$$

This Poisson distribution contains the estimable parameter of interest, μ, which is equal to the mean population size. This then gives the Poisson abundance Royle-Nichols model used in MacKenzie et al. (2006) and Royle (2006). Note that there is flexibility over $f(n)$ as there are alternative choices for mixing distributions. A regularly used two-parameter mixing distribution is the negative binomial Royle-Nichols model. Inference from using different mixing distributions can be seen in Stanley and Royle (2005, Tables 2 and 3).

It is worth mentioning the problems with *model* identifiability of occupancy parameters where different mixture functions give different estimates of occupancy. This can be seen in Royle (2006, Table 3) where they consider a variety of occupancy models with different mixture functions, and the resulting estimates of occupancy range between 0.51 and 0.91. While it is true that only one of those models estimates abundance, they all give different estimates for occupancy. As Royle (2006) states, "These results suggest that while identifiability of φ is an important problem to be aware of, it may not inhibit inference." Therefore care must be taken in selecting a suitable mixture function to use for inference in this model. Note that this is the consideration of model identifiability, rather than parameter identifiability which is what we are exploring in this thesis. We only consider using a Poisson mixing function in this thesis

for parameter redundancy analysis.

To obtain exhaustive summary terms from the likelihood in Equation (6.14), we need to specify some upper limit, N, for the summation, to avoid a likelihood which contains a infinite amount of terms. The idea in practice of this is that as $f(n) \approx 0$ when n is large, the addition of terms involving $f(n)$ for $n > N$, for some given N, is of negligible value. This function depends on the assumptions about n and whether low values of n are realistic or likely. We then obtain the likelihood for the Royle-Nichols model using a Poisson mixing function as

$$L(w) = \sum_{n=0}^{N} \binom{k}{w} \{1-(1-r)^n\}^w \{(1-r)^n\}^{k-w} \frac{e^{-\mu}\mu^n}{n!}. \tag{6.16}$$

Estimates of abundance are usually obtained from

$$Pr(n|w) = \frac{Pr(w|n) \cdot Pr(n)}{\sum_{n=1}^{N} Pr(w|n) \cdot Pr(n)}, \tag{6.17}$$

by the use of Bayes Theorem, where $Pr(n) = f(n)$ and $Pr(w|n)$ are binomial probabilities from the data as seen in Equation (6.13). While using this formula may give poor estimates of n at a local scale, average or total abundances from the data will be estimated well from this equation. Discussion of this can be seen in Royle and Nichols (2003). We can also obtain an estimate of occupancy, $\hat{\varphi}_0^{[1]}$, easily using this Royle-Nichols model. An estimate of occupancy can be simply formed as

$$\begin{aligned} \hat{\varphi}_0^{[1]} &= 1 - f(0) \\ &= 1 - e^{-\hat{\mu}}, \end{aligned}$$

using a Poisson mixing function in the model.

6.4.2 Adding group- and/or survey-dependency in the Royle-Nichols Poisson model

As a common theme in this thesis, we are interested in exploring what happens to a model's parameter redundancy when you change parameter dependencies in the model. In the Royle-Nichols model, we have two sets of parameters of interest which we can vary. The first parameter we can vary is the individual detection parameter r. We can let this parameter be group-dependent and/or survey-dependent by now denoting this parameter as $r_{s,j}$ during survey j for group s. If this parameter is not

group- or survey-dependent, then it is constant and the situation is equivalent to the previous section where there was just one parameter, r, regardless of survey or group. The Poisson mixing distribution has the parameter μ which could be constant for all groups, or be group-dependent, denoted as μ_s for group s. This means we can rewrite the likelihood with the most complex model with full group- and survey-dependence, conditional on observing n individuals, as

$$L = \sum_{n=0}^{N} \prod_{s=1}^{S} \prod_{j=1}^{k} \left\{1 - (1 - r_{s,j})^n\right\}^{\delta_{s,j}} \left\{(1 - r_{s,j})^n\right\}^{1-\delta_{s,j}} \frac{e^{-\mu_s}\mu_s^n}{n!}, \tag{6.18}$$

where $\delta_{s,j} = 1$ when there is a detection in survey j for group s, and $\delta_{s,j} = 0$ when there is *no* detection in survey j for group s, for all $j = 1, \ldots, k$ and $s = 1, \ldots, S$.

To generate the exhaustive summary terms we will use in this section, consider letting $f_{s,n} = e^{-\mu_s}\mu_s^n/n!$ which is the mixing function for a single n value of group s. It can be shown that a single occupancy-history, h_X, where n individuals are present in group s has the probability

$$Pr(h_X) = \prod_{j=1}^{k} \left[\left\{1 - (1 - r_{s,j})^n\right\}^{\delta_{s,j}} \left\{(1 - r_{s,j})^n\right\}^{1-\delta_{s,j}}\right] f_{s,n}. \tag{6.19}$$

This is equivalent to obtaining the joint probability of the observed data given from the $\delta_{s,j}$ values as well the number of individuals n. We show some example Royle-Nichols occupancy-history probabilities below.

Example 6.18: Consider the occupancy-history 011 for a study of $k = 3$ surveys. If we consider full group- and survey-dependence on both individual detection probabilities and mixture function probabilities, then the probability of the occupancy-history is equal to

$$Pr(h_X) = \left\{1 - r_{s,1}\right\}^n \left\{1 - (1 - r_{s,2})^n\right\} \left\{1 - (1 - r_{s,3})^n\right\} f_{s,n},$$

for any given number of individuals n for group s. Observe that n is not actually a parameter in the model but is one which can be given, and that $f_{s,n}$ represents the Poisson mixing function $f_{s,n} = e^{-\mu_s}\mu_s^n/n!$. □

We can consider missing observations in a single Royle-Nichols occupancy-history too. For an observation which has not been recorded that observation is effectively ignored. For example if the first survey was a missing observation, then there would

be no $r_{s,1}$ parameter given for that occupancy-history in its exhaustive summary term. This changes the probability of a single occupancy-history to be equal to

$$Pr(h_X) = \prod_{j=1}^{k} \left[\gamma_{s,j} \left\{ 1 - (1 - r_{s,j})^n \right\}^{\delta_{s,j}} \left\{ (1 - r_{s,j})^n \right\}^{1-\delta_{s,j}} + (1 - \gamma_{s,j}) \right] f_{s,n}, \quad (6.20)$$

where

$$\gamma_{s,j} = \begin{cases} 1 & \text{when } \delta_{s,j} \geq 0 \text{ (i.e. it not a missing observation),} \\ 0 & \text{when } \delta_{s,j} = -1 \text{ (i.e. it } \textit{is} \text{ a missing observation).} \end{cases}$$

Example 6.19: Consider the occupancy-history 0_11 for a study of $k = 4$ surveys. If we consider full group- and survey-dependence on both parameters as well as survey-dependent individual detection parameters, then the probability of the occupancy-history is equal to

$$Pr(h_X) = \{1 - r_{s,1}\}^n \{1 - (1 - r_{s,3})^n\} \{1 - (1 - r_{s,4})^n\} f_{s,n},$$

for any given number of n individual for group s. We see here that the second observation is effectively ignored so there is no $r_{s,2}$ parameter in the occupancy-history probability. □

6.4.3 Extrinsic parameter redundancy in the Royle-Nichols Poisson model

We can now see an example of the `Maple` code used to generate exhaustive summary terms from binary code occupancy-histories in Example 6.20. This code can be used to examine the extrinsic parameter redundancy of the model as for previous methods shown in Section 2.2.

Example 6.20: The Royle-Nichols model using a simulated data set: [See electronic appendix `example6.20.mw`] Consider a study where there are $k = 4$ surveys and we observe the following occupancy-histories: 1111, 1100, 1011, 1_0, 011_, 0100, 0010, 0000. The inputs for the `Maple` procedure `occmodroylenicholshist` require the vector **S2** as the column vector denoting which group each occupancy-history is in as

$$\mathbf{S2} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T,$$

and $\mathbf{T}$ as the matrix of all occupancy-histories as

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The maximum number of individuals which are present at the study location, N, also needs to be defined in this procedure. We set $N = 2$ in this example to reduce the complexity in these calculations but we can consider a model which has a much larger value of N. Say we use a Poisson mixture with no group-dependence so there are n unknown parameters, $f_{c,n}$, which represent the mixing functions in the model. This in effect means there is a single parameter for the Poisson mixture distribution, μ, but the mixing distribution is further dependent on the value of n taken as $f(c, n) = e^{-\mu}\mu^n/n!$. This means that there is a constant μ for all groups in the Poisson mixture distribution. We let the individual species detection probabilities be only survey-dependent so we have $r_{c,j}$ for $j = 1, \ldots, 4$. The exhaustive summary and parameter redundancy for this model can then be obtained by the code below.

```
S2 := Matrix(8,1,1):
# a 1 by 8 column vector with all entries as 1: Observe that
#  this vector still needs to be defined even if there is no
#  parameters which are group-dependent.
T := < <1|1|1|1>, <1|1|0|0>, <1|0|1|1>, <1|-1|-1|0>,
       <0|1|1|-1>, <0|1|0|0>, <0|0|1|0>, <0|0|0|0> >;
P := occmodroylenicholshist(3,1,2,S2,T);
# Inputs: (y,z,N,S2,T);
# y = individual detection probability;
# z = mixture parameter dependency;
# for y: '1'=constant, '2'=group-dep, '3'=survey-dep,
# '4'=group+survey-dep;
# for z: '1'=constant, '2'=group-dep;
# N = number of maximum individuals which could be present;
# S2 = column vector determining the group of each occupancy-history;
```

```
# T = list of all occupancy-histories.
kappa := Matvec(P);
theta := parsproc(kappa);
DD := Dmat(logvector(kappa),theta):
hybrid := Formnum2(DD,theta);
```

Let $\mathbf{P}$ be the matrix which contains all of the possible occupancy-history probabilities given in the data set which are dependent on the different values of n possible. Each row in $\mathbf{P}$ gives a different occupancy-history from the data, and each column j gives a different value of n. For example as $N = 2$, the first column would give the probabilities of the different occupancy-histories from the data when $n = 0$. The second column would give the probabilities of the occupancy-histories when $n = 1$, and the final column would give the probabilities of the occupancy-histories when $n = 2$. This matrix is given below

$$\mathbf{P} = \begin{bmatrix} 0 & r_{c,1}r_{c,2}r_{c,3}r_{c,4}f_{c,1} & \cdots \\ 0 & r_{c,1}r_{c,2}(1-r_{c,3})(1-r_{c,4})f_{c,1} & \cdots \\ 0 & r_{c,1}(1-r_{c,2})r_{c,3}r_{c,4}f_{c,1} & \cdots \\ 0 & r_{c,1}(1-r_{c,4})f_{c,1} & \cdots \\ 0 & (1-r_{c,1})r_{c,2}r_{c,3}f_{c,1} & \cdots \\ 0 & (1-r_{c,1})r_{c,2}(1-r_{c,3})(1-r_{c,4})f_{c,1} & \cdots \\ 0 & (1-r_{c,1})(1-r_{c,2})r_{c,3}(1-r_{c,4})f_{c,1} & \cdots \\ f_{c,0} & (1-r_{c,1})(1-r_{c,2})(1-r_{c,3})(1-r_{c,4})f_{c,1} & \cdots \end{bmatrix}$$

$$\begin{bmatrix} \cdots & \left\{1-(1-r_{c,1})^2\right\}\left\{1-(1-r_{c,2})^2\right\}\left\{1-(1-r_{c,3})^2\right\}\left\{1-(1-r_{c,4})^2\right\}f_{c,2} \\ \cdots & \left\{1-(1-r_{c,1})^2\right\}\left\{1-(1-r_{c,2})^2\right\}(1-r_{c,3})^2(1-r_{c,4})^2 f_{c,2} \\ \cdots & \left\{1-(1-r_{c,1})^2\right\}(1-r_{c,2})^2\left\{1-(1-r_{c,3})^2\right\}\left\{1-(1-r_{c,4})^2\right\}f_{c,2} \\ \cdots & \left\{1-(1-r_{c,1})^2\right\}(1-r_{c,4})^2 f_{c,2} \\ \cdots & (1-r_{c,1})^2\left\{1-(1-r_{c,2})^2\right\}\left\{1-(1-r_{c,3})^2\right\}f_{c,2} \\ \cdots & (1-r_{c,1})^2\left\{1-(1-r_{c,2})^2\right\}(1-r_{c,3})^2(1-r_{c,4})^2 f_{c,2} \\ \cdots & (1-r_{c,1})^2(1-r_{c,2})^2\left\{1-(1-r_{c,3})^2\right\}(1-r_{c,4})^2 f_{c,2} \\ \cdots & (1-r_{c,1})^2(1-r_{c,2})^2(1-r_{c,3})^2(1-r_{c,4})^2 f_{c,2} \end{bmatrix},$$

where $r_{c,j}$ denotes the individual detection probability in survey j and $f_{c,n}$ denotes the mixing Poisson parameter given there are n individuals in the study location. This mixing Poisson parameter $f_{c,n}$ is equivalent to obtaining μ in the Poisson mixture $f_{c,n} = e^{-\mu}\mu^n/n!$. The non-zero terms in $\mathbf{P}$ can then give an exhaustive summary for

this model as

$$\boldsymbol{\kappa} = \begin{bmatrix} r_{c,1}r_{c,2}r_{c,3}r_{c,4}f_{c,1} \\ \left\{1-(1-r_{c,1})^2\right\}\left\{1-(1-r_{c,2})^2\right\}\left\{1-(1-r_{c,3})^2\right\}\left\{1-(1-r_{c,4})^2\right\}f_{c,2} \\ r_{c,1}r_{c,2}(1-r_{c,3})(1-r_{c,4})f_{c,1} \\ \left\{1-(1-r_{c,1})^2\right\}\left\{1-(1-r_{c,2})^2\right\}(1-r_{c,3})^2(1-r_{c,4})^2f_{c,2} \\ r_{c,1}(1-r_{c,2})r_{c,3}r_{c,4}f_{c,1} \\ \vdots \\ (1-r_{c,1})^2(1-r_{c,2})^2\left\{1-(1-r_{c,3})^2\right\}(1-r_{c,4})^2f_{c,2} \\ f_{c,0} \\ (1-r_{c,1})(1-r_{c,2})(1-r_{c,3})(1-r_{c,4})f_{c,1} \\ (1-r_{c,1})^2(1-r_{c,2})^2(1-r_{c,3})^2(1-r_{c,4})^2f_{c,2} \end{bmatrix}.$$

The parameter set is given by $\boldsymbol{\theta} = [f_{c,0},\ f_{c,1},\ f_{c,2},\ r_{c,1},\ r_{c,2},\ r_{c,3},\ r_{c,4}]^T$. There are seven parameters in the model, four for the individual detection probabilities for each survey j, and three mixing parameters for each different $n = 0, 1, 2$ value as $N = 2$. The derivative matrix is not given here to save space but can be found in the `Maple` file `example6.20.mw`. We use the hybrid symbolic-numerical method of Section 2.5 to find the model rank due to computational complexity of the derivative matrix. This derivative matrix has the model rank and parameter deficiency as

$\mathbf{r} := \mathbf{7}, \mathbf{d} := \mathbf{0}$.

This means that this data set for the Royle-Nichols has no parameter redundancy and all the parameters in the model can be estimated.

We demonstrate how this procedure can be used with real data by using data sets on woodthrush and catbirds, taken from the North American Breeding Bird Survey (Robbins et al., 1986). These are two data sets which have regularly been used in the `unmarked` package for the statistical software package `R`, as shown in Fiske and Chandler (2011) and Fiske et al. (2014). These data sets have also previously been used to estimate population sizes using Poisson mixture distributions in Royle (2004b). Both data sets have 50 individual occupancy-histories observed over 11 surveys. The woodthrush data set is richer than the catbird data set as it contains a total of 206 observed detections during the 50 occupancy-histories, compared to the catbird data set which only has 51 observed detections. We arbitrarily divide this data set into 5 different groups each containing 10 different occupancy-histories so we can also view

Table 6.5: Parameter redundancies for Royle-Nichols models for the woodthrush and catbird data sets from Fiske et al. (2014)

	Woodthrush data set		Catbirds data set	
Model	Model rank	Deficiency	Model rank	Deficiency
$r(\cdot)\ \mu(\cdot)$	5	0	5	0
$r(\cdot)\ \mu(s)$	18	0	21	0
$r(s)\ \mu(\cdot)$	9	0	9	0
$r(s)\ \mu(s)$	22	0	25	0
$r(j)\ \mu(\cdot)$	15	0	15	0
$r(j)\ \mu(s)$	28	0	31	0
$r(s,j)\ \mu(\cdot)$	54	5	39	20
$r(s,j)\ \mu(s)$	64	8	51	24

group-dependency in this data set. In the model description, we denote the detection parameters as $r(\cdot)$ and the mixing parameters as $\mu(\cdot)$, with $(\cdot)$ denoting that the parameters are constant, (s) denoting that the parameters are group-dependent, and (j) denoting that the parameters are survey-dependent for the detection probabilities only. We use the `Maple` code in the electronic appendix file `example6.20.mw` to obtain the model rank and parameter deficiency for these data sets given in Table 6.5.

We can see from Table 6.5 that there is a reduction in the number of estimable parameters in the final two rows of the table. However, most models are full rank, so that all parameters can be estimated in those Royle-Nichols occupancy models. □

6.5 Discussion and further occupancy models

This chapter has gone beyond the work done in Chapter 5 and explored different model possibilities for occupancy studies. We have considered a general m-state model, though in practice a species will only have three or four different observational states which the animal could be in. We have developed `Maple` code to obtain individual occupancy-history probabilities and have used them to examine extrinsic parameter deficiencies in an example containing green frogs in Maryland in Section 6.1.5 from MacKenzie et al. (2009). Some of the results in Table 6.3 showed that including the use of state-dependent parameters can *decrease* parameter deficiency by un-confounding some of the parameters in some cases.

In Section 6.3 a two-species interaction model was considered as an extension of

the original occupancy model. This allowed the occupancy and detection parameters to be estimated dependent on an interaction between two different species. For example, the presence of species A could affect the presence of species B or vice-versa. We used `Maple` code to find the extrinsic parameter redundancies for the Yellowstone and Grand Teton National Park data example from Gould et al. (2012) and examined interactions between Columbian spotted frogs, boreal chorus frogs and tiger salamanders. The Royle-Nichols model was then considered by exploiting the link between animal abundance and probability of species occupation. We explored this model and show some extrinsic parameter redundancy results using `Maple` on data sets of woodthrush and catbirds from Fiske et al. (2014).

The analysis of these occupancy models is important as the use of these models is increasing. There is still future work to be done on this subject and we list some of the extensions that have not been considered in this thesis:

- The original occupancy model deals with the fact that non-detection does not imply occupancy, in other words taking into account false-negative errors. Incorporating false-positive errors which consider incorrect detections say by an incorrect species detection by the observer, is discussed in Royle and Link (2006), Green et al. (2011) and Miller et al. (2011).

- We can view different habitat states rather than occupancy states exploring how species prefer particular habitats. MacKenzie et al. (2011) consider this and obtain estimates of transition occupancy probabilities of species going between habitat states as an alternative to different occupancy states between seasons. A possible model with two habitat states could be based on the two-species interaction model of Section 6.3.

- We discretise each detection in our models but an alternative model can be fitted so each observation has a continuous time of detection. This can be seen in Hines et al. (2010) and Guillera-Arroita et al. (2011) where a multinomial likelihood can still be extracted from the observations recorded.

- We can view occupancy modelling to obtain estimates of species richness where we have imperfect detection, such as in Dorazio and Royle (2005) as well as Dorazio et al. (2006).

- We can consider the data in the form of detection counts for a survey instead of recording it was detected or not detected at the survey. This is seen in a variety

of contexts such as in Royle (2004a), Kéry et al. (2009), Yamaura et al. (2011) and Guillera-Arroita et al. (2012).

- There has also been recent work done on exploring the connection between occupancy models and multi-event capture-recapture models, as seen in Pradel, 2005 and Gimenez et al. (2014) by considering hidden Markov models, where they provide a direct reparameterisation between the two models.

Chapter 7

Conclusions and Future Work

7.1 Discussion

In this thesis, we have examined parameter redundancy for a variety of ecological models. Statistical ecology itself is an ever increasing area of research as more people nowadays require information on how ecological ecosystems behave and interact in the world. This information influences government and business environmental policies so that they can best protect these ecosystems. More complex ecological models are used to analyse such ecosystems due to the increase of more advanced technologies for both collection and interpretation of data, as shown in King (2014). As these government and business policies require parameter estimates such as survival or occupancy probabilities, parameter redundancy identifies whether statistical models can actually obtain these parameter estimates in principle. A model which does not generate the full estimation of the model parameters is not ideal, as obtaining these values is what is required from the analysis using the model. In such cases, parameter constraints are then used to obtain inference, as regularly used in linear modelling in other statistical fields. This thesis explores these issues and provides lists of ecological models which are not parameter redundant so that these parameters can be estimated.

We have two purposes for this work: The first purpose is to identify models that are intrinsically parameter redundant. It would not be wise to use a model for an analysis of a data set where there is some confounding of the parameters, meaning you cannot achieve the parameter estimates that are required. We show which a variety of models which are not parameter redundant and indicate what parameter constraints can be used so that the model is not intrinsically parameter redundant. The second purpose is to give an indication as to how rich a data set needs to be for a particular model

to be able to estimate every parameter in the model. We use example data sets to show typical extrinsic parameter redundancies. The results of these analyses show that the data can be far from 'perfect', where not all of the possibilities are observed during a study, yet all the parameters can still be estimated for some of the simpler models.

A wide variety of models and applications have been considered in this thesis. Chapter 2 examines mark-recovery models where we mark animals and then attempt to recover these marked animals dead at future time points. The objective of this model is to find estimates of animal survival between time points in the study. This is also the case in capture-recapture models of Chapter 3 where we estimate survival probabilities by the use of live recaptures rather than dead recoveries. This is further extended in Chapter 4 where live recaptures and dead recoveries are both considered in the same model. Some of the models in this thesis show that even though these seem attractive models to use, they can have issues with parameter redundancy even for relatively simple models, such as the fully time-dependent capture-recapture model. We consider a range of models where we change the parameter dependencies in the model, and age- and time-dependency are common parameter dependencies we examine. The majority of capture-recapture and capture-recapture-recovery models are not intrinsically parameter redundant so the parameter redundancy in these cases usually came from sparse data sets rather than the model itself. Examples of data sets explored for extrinsic parameter redundancy includes data on dippers from Marzolin (1988) in capture-recapture modelling and data on cormorants from Hènaux et al. (2007) in capture-recapture-recovery modelling.

We then proceeded to examine a different ecological model which estimates species occupancy rates rather than survival probabilities. Chapter 5 presents how these occupancy models can be used to generate estimates of species occupancy using a robust design with more than one survey per season for multiple seasons. It is shown that the occupancy models of Chapter 5 are intrinsically not parameter redundant given that each season has at least two surveys during each season, meaning this model in theory is an ideal one to use as all of the model parameters can be estimated. However, what is usually the case in occupancy studies is that you record data which are far from perfect so the analysis of sparse data sets becomes more crucial to consider in these models. To show the problems of sparse data, we view data on house finches from MacKenzie et al. (2006) in Table 5.1, as well as considering group-dependent data in Table 5.2 on amphibian breeding in Yellowstone and Grand Teton national parks in USA, given in Gould et al. (2012). Missing observations in the data also becomes more of an

issue in occupancy studies, and these extrinsic parameter redundancy results are then considered by also including partial occupancy-histories. This increases the number of occupancy-histories in the data and reduces the extrinsic parameter redundancies as seen in Table 5.3. A brief analysis on approximate sample sizes was also explored in Section 5.9 and shows that even more data are required when transition and detection probabilities in a study are low. These ideas were extended further in Chapter 6 where a collection of alternative models were examined for extrinsic parameter redundancy using a range of data sets. These included the addition of multiple states in Section 6.1, the two-species interaction model in Section 6.3, and the Royle-Nichols model in Section 6.4.

7.2 Further work

There is still much work to be done on parameter redundancy in the future. We list a few ways parameter redundancy analysis can be developed:

- While a range of models have been examined for parameter redundancy as seen in Tables 1.1a and 1.1b, it is certainly not exhaustive and there are many more models to consider. Distance sampling in ecology (such as reviewed in Thomas et al., 2010) is one such wide area that requires parameter redundancy analysis. Other models include spatially explicit capture-recapture models (such as in Borchers, 2012), Jolly-Seber tag loss models (such as in Cowen and Schwarz, 2006), multi-event modelling (such as in Pradel, 2005) and further occupancy models extensions as discussed in Section 6.5.

- We would like to deal with some potential problems in the estimation of parameter combinations when the exhaustive summary is complex. Currently the `Maple` code for generating the partial differential equations and then the estimable parameter combinations fails to work due to a lack of computational memory. This is something that could be considered and ways of solving these PDE's with less computational complexity could be examined. The PDE's can be solved numerically so the values of $\boldsymbol{\alpha}$ can be obtained, so we can conclude which original parameters can be explicitly estimated, but we have no method of obtaining the other estimable parameter combinations from the model. Recent work in identifiability analysis has also considered this issue (Eisenberg and Hayashi, 2014).

- Identifiability issues can be further be considered in a Bayesian framework as shown in Gelfand and Sahu (1999), Garrett and Zeger (2000) and Gimenez et al.

(2009). They define weak identifiability as when the data give little information about certain parameters in the model, such as how parameter redundancy affects the parameter estimation in classical techniques. Similarly to the classical approach, parameter estimation conclusions using the Bayesian approach can be misleading or incorrect if there is weak identifiability. We only used classical approaches in this thesis as they are easier to code than MCMC methods. This area of research with the use of priors to help inference is an alternative to using model constraints and is something that could be explored in more depth.

- Finally, a piece of software could be programmed which makes the calculation of a model's parameter redundancy easy and user-friendly. This could come in the form of a stand-alone program and would give users step-by-step instructions on how to input matrices which generates final parameter redundancy results without having to think about the codes behind the procedure. Similar ideas have been considered in the past regarding identifiability analysis in the software package `DAISY`, see Bellu et al. (2007) for further details. Currently the program `Maple` needs to be used to do the calculations as given in this thesis, and this process could be made easier for users without `Maple` knowledge.

Appendix A

Common Maple Procedures

This Appendix provides the basic `Maple` procedures used in this thesis. We include these basic procedures which allow us to examine the parameter redundancy and give the estimable parameter combinations of a particular model. These procedures are mostly the work of previous authors including Catchpole et al. (2002), Cole et al. (2010) and Choquet and Cole (2012) and we list these basic procedure below:

1. `Matvec`: Converts a data matrix into its vector of non-zero terms.
2. `logvector`: Converts a data matrix into its log-vector form of the exhaustive summary $\boldsymbol{\kappa}$(see Section 2.4).
3. `parsproc`: Generates a list of all the parameters in the model.
4. `Dmat`: Forms the model derivative matrix by differentiating the exhaustive summary by its parameters.
5. `Formnum`: Performs the hybrid symbolic-numerical method of Choquet and Cole (2012) to find the model's rank and parameter deficiency along with $\boldsymbol{\alpha}$(see Section 2.5).
6. `Formnum2`: The same hybrid symbolic-numerical method as `Formnum` to find the model's rank and parameter deficiency but without the representation of $\boldsymbol{\alpha}$ in the output (see Section 2.5).
7. `Estpars`: Finds all of the estimable parameter combinations for parameter redundant models.

The longer procedures developed as part of the thesis can be found in the online supplementary `Maple` files for the different examples during the thesis. This includes the `Maple` codes that obtain a model's simpler intrinsic exhaustive summary (such as

`ringmod` from Chapter 2 or `capmod` from Chapter 3), or the codes that turn animal histories into their multinomial probabilities to explore extrinsic parameter redundancy (such as `caprecaphistories` from Chapter 3 or `occmodseasonshist` from Chapter 5).

1. **Matvec** := proc(**P**)
local *sizekappa*, i, j, $\boldsymbol{\kappa}$, *kappaindex*;
sizekappa := 0:
for i to Dimension(**P**)[1] do
for j to Dimension(**P**)[2] do
if $\mathbf{P}[i,j] \neq 0$ then *sizekappa* := *sizekappa* + 1
end if:
end do: end do:
$\boldsymbol{\kappa}$:= Vector(*sizekappa*):
kappaindex := 0:
for i to Dimension(**P**)[1] do
for j to Dimension(**P**)[2] do
if $\mathbf{P}[i,j] \neq 0$ then *kappaindex* := *kappaindex* + 1:
$\boldsymbol{\kappa}$[*kappaindex*] := $\mathbf{P}[i,j]$:
end if:
end do: end do:
$\boldsymbol{\kappa}$
end proc:

2. **logvector** := proc($\boldsymbol{\kappa}$)
local i, **lnkappa**;
lnkappa := Vector($\boldsymbol{\kappa}$):
for i to Dimension($\boldsymbol{\kappa}$) do **lnkappa**$[i]$:= ln($\boldsymbol{\kappa}[i]$):
end do:
lnkappa
end proc:

3. **parsproc** := proc($\boldsymbol{\kappa}$)
$\langle$(seq(op(n, indets($\boldsymbol{\kappa}$)), $n = 1$ .. nops(indets($\boldsymbol{\kappa}$)))$\rangle$;
end proc:

4. **Dmat** := proc($\boldsymbol{\kappa}$,$\boldsymbol{\theta}$)
local **DD1**, i, j;

DD1 := Matrix(1 .. Dimension($\boldsymbol{\theta}$), 1 .. Dimension($\boldsymbol{\kappa}$)):
for i to Dimension($\boldsymbol{\theta}$) do
for j to Dimension($\boldsymbol{\kappa}$) do **DD1**$[i, j]$:= diff($\boldsymbol{\kappa}[j]$, $\boldsymbol{\theta}[i]$)
end do: end do:
DD1
end proc:

5. **Formnum** := proc(**D1**,$\boldsymbol{\theta}$)
local j, **results**, *numpars*, **D1rand**;
results := Matrix(5,3):
for j from 1 to 5 do *numpars* := seq$\left(\boldsymbol{\theta}[i] = \text{evalf}\left(\frac{rand(\,)}{1000000000000}\right),\ i = 1 \text{ .. Dimension}(\boldsymbol{\theta})\right)$:
D1rand := eval(**D1**, *numpars*);
results[j,1] := Rank(**D1rand**);
results[j,2] := Dimension($\boldsymbol{\theta}$) - Rank(**D1rand**);
results[j,3] := NullSpace(Transpose(**D1rand**));
end do:
results
end proc:

6. **Formnum2** := proc(**D1**,$\boldsymbol{\theta}$)
local j, **results**, *numpars*, **D1rand**;
results := Matrix(5,2):
for j from 1 to 5 do *numpars* := seq$\left(\boldsymbol{\theta}[i] = \text{evalf}\left(\frac{rand(\,)}{1000000000000}\right),\ i = 1 \text{ .. Dimension}(\boldsymbol{\theta})\right)$:
D1rand := eval(**D1**, *numpars*);
results[j,1] := Rank(**D1rand**);
results[j,2] := Dimension($\boldsymbol{\theta}$) - Rank(**D1rand**);
end do:
results
end proc:

7. **Estpars** := proc(**D1**,$\boldsymbol{\theta}$)
local i, r, d, **alphapre**, $\boldsymbol{\alpha}$, **PDE**, *FF*, **ans**;
r := Rank(**D1**);
d := Dimension($\boldsymbol{\theta}$) - r;
alphapre := NullSpace(Transpose(**D1**));
$\boldsymbol{\alpha}$:= Matrix(d, Dimension($\boldsymbol{\theta}$));
PDE := Vector(d);

FF := f(seq($\boldsymbol{\theta}$[i], i = 1 .. Dimension($\boldsymbol{\theta}$)));
for i to d do $\boldsymbol{\alpha}$[i, 1 .. Dimension($\boldsymbol{\theta}$)] := **alphapre**[i];
PDE[i] := add((diff(FF,$\boldsymbol{\theta}$[j]))·$\boldsymbol{\alpha}$[i, j], j = 1 .. Dimension($\boldsymbol{\theta}$))
end do;
ans := pdsolve(seq(**PDE**[i] = 0, i = 1 .. d));
⟨$\boldsymbol{\alpha}$, **PDE**, **ans**⟩
end proc:

Appendix B

Reparameterisation Proofs

B.1 Proof of the multiple capture-recapture simpler exhaustive summary theorem

The proof of Theorem 3.2 is also given in the supplementary material in Hubbard et al. (2014). The theorem states that a simpler exhaustive summary for the capture-recapture model consists of the terms $s_{i,j} = \phi_{i,j}p_{i+1,j+1}$ (for $i = 1, \ldots, n_2$ and $j = i, \ldots, \min(n_1 + i - 1, n_2)$) and $t_{i,j} = \phi_{i,j}(1 - p_{i+1,j+1})$ (for $i = 1, \ldots, n_2 - 1$ and $j = i, \ldots, \min(n_1 + i - 1, n_2 - 1)$). The proof of Theorem 3.2 is split into three parts:

- In part one, we show that the original exhaustive summary consisting of the capture-histories can be reparameterised in terms of $[\mathbf{s}, \mathbf{t}]$, utilising the reparameterisation theorem of Theorem 2.3.
- In part two, we create a new exhaustive summary, denoted as $\boldsymbol{\kappa}_{uvw}$, utilising the reparameterisation theorem of Theorem 2.3. This is created so that the extension theorem of Theorem 2.2 can be applied in order for results to be extended to any dimension.
- In part three, we show that the reparameterisation $[\mathbf{s}, \mathbf{t}]$ is an exhaustive summary, again utilising the reparameterisation theorem of Theorem 2.3.

We assume that none of the parameters are on boundary values, so that our parameter space is restricted to $0 < \phi_{i,j} < 1$ and $0 < p_{i,j} < 1$ for all values of i and j for this theorem to apply. Theorem 2.3 states that if the derivative matrix $[\partial\boldsymbol{\kappa}(\boldsymbol{\kappa}_{\text{new}})/\partial\boldsymbol{\kappa}_{\text{new}}]$ is full rank then $\boldsymbol{\kappa}_{\text{new}}$ is a new exhaustive summary.

Part one:

We show that any capture-history can be expressed in terms of parameters $s_{i,j}$ and $t_{i,j}$. The probability of a particular capture-history, h_X, with individual capture observations δ_k at time k, where an animal is first recaptured at time a and last recaptured at time b, is given as

$$Pr(h_X) = \left[\prod_{k=a+1}^{b} \phi_{k-a,k-1}\left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\}\right] \chi_{b-a+1,b}.$$

These probabilities can be reparameterised in terms of $s_{i,j} = \phi_{i,j}p_{i+1,j+1}$ and $t_{i,j} = \phi_{i,j}(1-p_{i+1,j+1})$ to give

$$Pr(h) = \prod_{k=a+1}^{b} \left(\delta_k s_{k-a+1,k} + (1-\delta_k)t_{k-a+1,k}\right) \chi_{b-a+1,b}.$$

The probability of an animal never being seen again, $\chi_{i,j}$, with $\chi_{i,n_2+1} = 1$, can be shown to be a function of $s_{i,j}$ and $t_{i,j}$, by first expanding $\chi_{i,j}$, to give

$$\begin{aligned}\chi_{i,j} \;=\; & (1-\phi_{i,j}) + \phi_{i,j}(1-p_{i+1,j+1}) \cdot [(1-\phi_{i+1,j+1}) + \\ & \phi_{i+1,j+1}(1-p_{i+2,j+2}) \cdot [(1-\phi_{i+2,j+2}) + \phi_{i+2,j+2} \cdot \\ & (1-p_{i+3,j+3}) \cdot [\cdots [((1-\phi_{n_2,n_2}) + \phi_{n_2,n_2}(1-p_{n_2+1,n_2+1})] \cdots]]].\end{aligned}$$

By noting that $(1-\phi_{i,j}) = (1-s_{i,j}-t_{i,j})$, we can write $\chi_{i,j}$ as

$$\begin{aligned}\chi_{i,j} \;=\; & (1-s_{i,j}-t_{i,j}) + t_{i,j}[(1-s_{i+1,j+1}-t_{i+1,j+1}) + t_{i+1,j+1} \cdot \\ & [(1-s_{i+2,i+2}-t_{i+2,j+2}) + t_{i+3,j+3}[\cdots [(1-s_{n_2,n_2})] \cdots]]].\end{aligned}$$

Therefore all capture-histories can be parameterised in terms of **s** and **t** only.

Part two:

We derive a new exhaustive summary which consists of the following terms:

- $u_j = \prod_{k=1}^{n_2-j+1} \phi_{k,k+j-1}p_{k+1,k+j}$,
  for all $j = 1, \ldots, n_2$;
- $v_{i,j} = \dfrac{(1-p_{i+1,j+1})}{p_{i+1,j+1}}$,
  for all $i = 1, \ldots, n_2 - 1$ and $j = i, \ldots, \min(i+n_1-1, n_2)$;

- and $w_{i,j} = \dfrac{\chi_{i+1,j+1}}{\left(\prod_{k=1}^{n_2-j} \phi_{k+i,k+j} p_{k+i+1,k+j+1}\right)}$,
  for all $i = 1, \ldots, n_2 - 1$ and $j = i, \ldots, \min(i + n_1 - 1 \leq n_2)$.

We can reparameterise the original exhaustive summary consisting of the capture-histories, when there are $n_1 = n_2 = 2$ years of marking and recapture, as

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(111) \\ Pr(101) \\ Pr(011) \\ Pr(110) \\ Pr(100) \\ Pr(010) \end{bmatrix} = \begin{bmatrix} \phi_{1,1} p_{2,2} \phi_{2,2} p_{3,3} \\ \phi_{1,1}(1 - p_{2,2}) \phi_{2,2} p_{3,3} \\ \phi_{1,2} p_{2,3} \\ \phi_{1,1} p_{2,2} \chi_{2,2} \\ \chi_{1,1} \\ \chi_{1,2} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_1 v_{1,1} \\ u_2 \\ u_1 w_{1,1} \\ \bar{u}_1 - u_1(v_{1,1} + w_{1,1}) \\ \bar{u}_2 \end{bmatrix}.$$

The reparameterisation is given as $\boldsymbol{\kappa}_{uvw} = [u_1, u_2, v_{1,1}, w_{1,1}]^T$ where the derivative matrix,

$$\left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\kappa}_{uvw}}\right] = \begin{bmatrix} 1 & v_{1,1} & 0 & w_{1,1} & -1 - v_{1,1} - w_{1,1} & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & u_1 & 0 & 0 & -u_1 & 0 \\ 0 & 0 & 0 & u_1 & -u_1 & 0 \end{bmatrix},$$

has full rank 4. A modified PLUR decomposition of $[\partial \boldsymbol{\kappa} / \partial \boldsymbol{\kappa}_{uvw}]$ shows that this is valid for all values of $u_1, u_2, v_{1,1}$ and $w_{1,1}$ as long as $u_1 = \phi_{1,1} p_{2,2} \phi_{2,2} p_{3,3}$ is non-zero. This only occurs at a boundary and the parameter space has already been restricted to exclude boundary values. Therefore by Theorem 2.3, when $n_1 = n_2 = 2$, $\boldsymbol{\kappa}_{uvw}$ is an alternative exhaustive summary for the model. Now consider extending the model firstly by adding another year of recapture so that $n_2 = 3$, while keeping $n_1 = 2$. The

original exhaustive summary is then

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(1111) \\ Pr(1011) \\ Pr(0111) \\ Pr(1100) \\ Pr(1000) \\ Pr(0101) \\ Pr(1110) \\ Pr(1010) \\ Pr(0110) \\ Pr(1101) \\ Pr(1001) \\ Pr(0100) \end{bmatrix} = \begin{bmatrix} \phi_{1,1}p_{2,2}\phi_{2,2}p_{3,3}\phi_{3,3}p_{4,4} \\ \phi_{1,1}\bar{p}_{2,2}\phi_{2,2}p_{3,3}\phi_{3,3}p_{4,4} \\ \phi_{1,2}p_{2,3}\phi_{2,3}p_{3,4} \\ \phi_{1,1}p_{2,2}\chi_{2,2} \\ \chi_{1,1} \\ \phi_{1,2}\bar{p}_{2,3}\phi_{2,3}p_{3,4} \\ \phi_{1,1}p_{2,2}\phi_{2,2}p_{3,3}\chi_{2,2}) \\ \phi_{1,1}\bar{p}_{2,2}\phi_{2,2}p_{3,3}\chi_{2,2} \\ \phi_{1,2}p_{2,3}\chi_{2,3} \\ \phi_{1,1}p_{2,2}\phi_{2,2}\bar{p}_{3,3}\phi_{3,3}p_{4,4} \\ \phi_{1,1}\bar{p}_{2,2}\phi_{2,2}\bar{p}_{3,3}\phi_{3,3}p_{4,4} \\ \bar{\phi}_{1,2} + \phi_{1,2}\bar{p}_{2,3}\chi_{2,3} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_1v_{1,1} \\ u_2 \\ u_1w_{1,1} \\ k_4 \\ u_2v_{1,2} \\ u_1w_{2,2} \\ u_1v_{1,1}w_{2,2} \\ u_2w_{1,2} \\ u_1v_{2,2} \\ u_1v_{1,1}v_{2,2} \\ \bar{u}_2 - u_2(v_{1,2} + w_{1,2}) \end{bmatrix},$$

where $k_4 = 1 - u_1 - u_1v_{1,1} - u_1w_{1,1} - u_1w_{2,2} - u_1v_{1,1}w_{2,2} - u_1v_{2,2} - u_1v_{1,1}v_{2,2}$ and $\bar{\phi} = (1-\phi)$. This uses the reparameterisation $\boldsymbol{\kappa}_{uvw} = [u_1, u_2, v_{1,1}, v_{1,2}, v_{2,2}, w_{1,1}, w_{1,2}, w_{2,2}]^T$. We now use the two-stage extension theorem of Theorem 3.1. The first stage involves the exhaustive summary terms

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} u_1 \\ u_1v_{1,1} \\ u_2 \\ u_1w_{1,1} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [u_1, u_2, v_{1,1}, w_{1,1}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_1}{\partial \boldsymbol{\theta}_1}\right] = \begin{bmatrix} 1 & v_{1,1} & 0 & w_{1,1} \\ 0 & 0 & 1 & 0 \\ 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & u_1 \end{bmatrix},$$

has full rank 4. The second stage examines the remaining exhaustive summary terms

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} k_4 \\ u_2 v_{1,2} \\ u_1 w_{2,2} \\ u_1 v_{1,1} w_{2,2} \\ u_2 w_{1,2} \\ u_1 v_{2,2} \\ u_1 v_{1,1} v_{2,2} \\ 1 - u_2 - u_2 v_{1,2} - u_2 w_{1,2} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_2 = [v_{1,2}, v_{2,2}, w_{1,2}, w_{2,2}]^T$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_2}{\partial \boldsymbol{\theta}_2}\right] = \begin{bmatrix} 0 & u_2 & 0 & 0 & 0 & 0 & 0 & -u_2 \\ -u_1 - u_1 v_{1,1} & 0 & 0 & 0 & 0 & u_1 & u_1 v_{1,1} & 0 \\ 0 & 0 & 0 & 0 & u_2 & 0 & 0 & -u_2 \\ -u_1 - u_1 v_{1,1} & 0 & u_1 & u_1 v_{1,1} & 0 & 0 & 0 & 0 \end{bmatrix},$$

has full rank 4. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of years of recapture. Adding a year of marking so $n_1 = 3$ while $n_2 = 3$ adds the following exhaustive summary terms

$$\boldsymbol{\kappa}_3 = \begin{bmatrix} Pr(0011) \\ Pr(0010) \end{bmatrix} = \begin{bmatrix} \phi_{1,3} p_{2,4} \\ 1 - \phi_{1,3} p_{2,4} \end{bmatrix} = \begin{bmatrix} u_3 \\ 1 - u_3 \end{bmatrix}.$$

As there is only one additional parameter in $\boldsymbol{\kappa}_3$, this is trivially full rank and means that the original model can be extended for a greater number years of marking. Therefore $\boldsymbol{\kappa}_{uvw}$ is an exhaustive summary for any dimension.

Part three:

This part involves checking whether the derivative matrix $[\partial \boldsymbol{\kappa}_{uvw}(\mathbf{s}, \mathbf{t}) / \partial (\mathbf{s}, \mathbf{t})]$ is full rank and then using the two-stage theorem of Theorem 3.1 to show it is always full rank for larger dimensions. Starting with $n_1 = n_2 = 2$ we can reparameterise $\boldsymbol{\kappa}_{uvw}$

in terms of $\mathbf{s}$ and $\mathbf{t}$ to get

$$\boldsymbol{\kappa}_{uvw}(\mathbf{s},\mathbf{t}) = \begin{bmatrix} u_1 \\ v_{1,1} \\ w_{1,1} \\ u_2 \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2} \\ t_{1,1}/s_{1,1} \\ (1-s_{2,2})/s_{2,2} \\ s_{1,2} \end{bmatrix}, \tag{B.1}$$

with parameter set $\boldsymbol{\kappa}_{st} = [s_{1,1}, s_{1,2}, s_{2,2}, t_{1,1}]^T$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_{uvw}(\mathbf{s},\mathbf{t})}{\partial \boldsymbol{\kappa}_{st}}\right] = \begin{bmatrix} s_{2,2} & -\frac{t_{1,1}}{s_{1,1}^2} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ s_{1,1} & 0 & -\frac{1}{s_{2,2}} - \frac{(1-s_{2,2})}{s_{2,2}^2} & 0 \\ 0 & \frac{1}{s_{1,1}} & 0 & 0 \end{bmatrix},$$

has full rank 4. A modified PLUR decomposition of $[\partial \boldsymbol{\kappa}_{uvw}(\mathbf{s},\mathbf{t})/\partial \boldsymbol{\kappa}_{st}]$ shows this the model remains full rank for any value of $s_{1,1}, s_{1,2}, s_{2,2}$ or $t_{1,1}$. Therefore $\boldsymbol{\kappa}_{st}$ is an exhaustive summary when $n_1 = n_2 = 2$. If we extend the model to add another year of recapture, the exhaustive summary becomes

$$\boldsymbol{\kappa}_{uvw}(\mathbf{s},\mathbf{t}) = \begin{bmatrix} u_1 \\ u_2 \\ v_{1,1} \\ v_{1,2} \\ v_{2,2} \\ w_{1,1} \\ w_{1,2} \\ w_{2,2} \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2}s_{3,3} \\ s_{1,2}s_{2,3} \\ t_{1,1}/s_{1,1} \\ t_{1,2}/s_{1,2} \\ t_{2,2}/s_{2,2} \\ (1-s_{3,3})/s_{3,3} \\ (1-s_{2,3})/s_{2,3} \\ \{(1-s_{2,2}-t_{2,2}) + t_{2,2}\,(1-s_{3,3})\}/s_{2,2}s_{3,3} \end{bmatrix},$$

with parameters $\boldsymbol{\kappa}_{st} = [s_{1,1}, s_{1,2}, s_{2,2}, s_{2,3}, s_{3,3}, t_{1,1}, t_{1,2}, t_{2,2}]^T$. Note that the terms $u_2, v_{2,2}$ and $w_{1,2}$ are identical to $u_1, v_{1,1}$ and $w_{1,1}$ respectively if $s_{1,1}$ is re-labelled as $s_{1,2}$, $s_{2,2}$ as $s_{2,3}$, and $t_{1,1}$ as $t_{1,2}$. This can then form the first stage of the two-stage extension theorem with

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} u_2 \\ v_{1,2} \\ w_{1,2} \end{bmatrix} = \begin{bmatrix} s_{1,2}s_{2,3} \\ t_{1,2}/s_{1,2} \\ (1-s_{2,3})/s_{2,3} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [s_{1,2}, s_{2,3}, t_{1,2}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_1}{\partial \boldsymbol{\theta}_1}\right] = \begin{bmatrix} s_{2,3} & -\frac{t_{1,2}}{s_{1,2}^2} & 0 \\ s_{1,2} & 0 & -\frac{1}{s_{2,3}} - \frac{(1-s_{2,3})}{s_{2,3}^2} \\ 0 & \frac{1}{s_{1,2}} & 0 \end{bmatrix},$$

has full rank 3. The second stage involves the terms

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} u_1 \\ v_{1,1} \\ v_{2,2} \\ w_{1,1} \\ w_{2,2} \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2}s_{3,3} \\ t_{1,1}/s_{1,1} \\ t_{2,2}/s_{2,2} \\ (1-s_{3,3})/s_{3,3} \\ \{(1-s_{2,2}-t_{2,2}) + t_{2,2}(1-s_{3,3})\}/s_{2,2}s_{3,3} \end{bmatrix},$$

with the parameter set $\boldsymbol{\theta}_2 = [s_{1,1}, s_{2,2}, s_{3,3}, t_{1,1}, t_{2,2}]^T$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_2}{\partial \boldsymbol{\theta}_2}\right] = \begin{bmatrix} s_{2,2}s_{3,3} & -\frac{t_{1,1}}{s_{1,1}^2} & 0 & 0 & 0 \\ s_{1,1}s_{3,3} & 0 & -\frac{t_{2,2}}{s_{2,2}^2} & 0 & -\frac{1}{s_{2,2}s_{3,3}} - \frac{\{(1-s_{2,2}-t_{2,2})+t_{2,2}(1-s_{3,3})\}}{s_{2,2}^2 s_{3,3}} \\ s_{1,1}s_{2,2} & 0 & 0 & -\frac{1}{s_{2,2}} - \frac{(1-s_{3,3})}{s_{2,2}^2} & -\frac{t_{2,2}}{s_{2,2}s_{3,3}} - \frac{\{(1-s_{2,2}-t_{2,2})+t_{2,2}(1-s_{3,3})\}}{s_{2,2}s_{3,3}^2} \\ 0 & \frac{1}{s_{1,1}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{s_{2,2}} & 0 & -\frac{1}{s_{2,2}} \end{bmatrix},$$

has full rank 5. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of years of recapture. Adding a year of marking so $n_1 = 3$ while $n_2 = 3$ adds only the exhaustive summary term $u_3 = s_{1,3}$. As there is only one additional parameter, this extension is trivially full rank and means that the original model can be extended for a greater number of years of marking. Therefore $\mathbf{s}$ and $\mathbf{t}$ form an exhaustive summary for any dimension. This then completes the proof of Theorem 3.2. □

Note that alternative exhaustive summaries could have been used here as well. This includes only considering the exhaustive summaries generated by the parameters $s_{i,j}$ and $s_{i,j} + t_{i,j} = \phi_{i,j}$, or alternatively $s_{i,j}$ and $s_{i,j}/(s_{i,j} + t_{i,j}) = p_{i,j}$. Both these alternative exhaustive summaries would obtain the same general intrinsic parameter redundancy result of Theorem 3.2 as shown by use of the reparameterisation theorem of Theorem 3.2.

B.2 Proof of the multiple capture-recapture-recovery simpler exhaustive summary theorem

The proof of Theorem 4.1 is also given in the supplementary material in Hubbard et al. (2014). The theorem states that a simpler exhaustive summary for the multiple capture-recapture-recovery model consists of the terms $s_{i,j} = \phi_{i,j}p_{i+1,j+1}$ (for $i = 1,\ldots,n_2$ and $j = i,\ldots,\min(n_1+i-1,n_2)$), $t_{i,j} = \phi_{i,j}(1-p_{i+1,j+1})$ (for $i = 1,\ldots,n_2-1$ and $j = i,\ldots,\min(n_1+i-1,n_2-1)$) and $r_{i,j} = (1-\phi_{i,j})\lambda_{i,j}$ (for all $i = 1,\ldots,n_2$ and $j = i,\ldots,\min(n_1+i-1,n_2)$). The proof of Theorem 4.1 is split into three parts:

- In part one, we show that the original exhaustive summary consisting of the life-histories can be reparameterised in terms of $[\mathbf{s},\mathbf{t},\mathbf{r}]$, utilising the reparameterisation theorem of Theorem 2.3.
- In part two, we create a new exhaustive summary, denoted as $\boldsymbol{\kappa}_{uvwx}$, utilising the reparameterisation theorem of Theorem 2.3. This is created so that the extension theorem of Theorem 2.2 can be applied in order for results to be extended to any dimension.
- In part three, we show that the reparameterisation $\boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r})$ is an exhaustive summary, again utilising the reparameterisation theorem of Theorem 2.3.

We assume that none of the parameters are on boundary values, so that our parameter space is restricted to $0 < \phi_{i,j} < 1$, $0 < p_{i,j} < 1$ and $0 < \lambda_{i,j} < 1$ for all values of i and j for this theorem to apply. Theorem 2.3 states that if the derivative matrix $[\partial\boldsymbol{\kappa}(\boldsymbol{\kappa}_{\text{new}})/\partial\boldsymbol{\kappa}_{\text{new}}]$ is full rank then $\boldsymbol{\kappa}_{\text{new}}$ is a new exhaustive summary.

Part one:

We show that any life-history can be expressed in terms of parameters $s_{i,j}$, $t_{i,j}$ and $r_{i,j}$. The probability of a particular life-history, h_X, with individual capture/recovery observations δ_k at time k, where an animal is first recaptured at time a and last recaptured or recovered at time b, is given as

$$Pr(h_X) = \begin{cases} \displaystyle\prod_{k=a+1}^{b} \phi_{k-a,k-1}\left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\}\chi_{b-a+1,b} & \text{if } \delta_b = 1, \\ \displaystyle\prod_{k=a+1}^{b-1} \phi_{k-a,k-1}\left\{\delta_k p_{k-a+1,k} + (1-\delta_k)(1-p_{k-a+1,k})\right\} & \text{if } \delta_b = 2. \\ \qquad\qquad \times(1-\phi_{b-a,b-1})\lambda_{b-a,b-1} \end{cases}$$

These probabilities can be reparameterised in terms of $s_{i,j} = \phi_{i,j}p_{i+1,j+1}$, $t_{i,j} = \phi_{i,j}(1-p_{i+1,j+1})$ and $r_{i,j} = (1-\phi_{i,j})\lambda_{i,j}$ to give

$$Pr(h) = \begin{cases} \prod_{k=a+1}^{b} \{\delta_k s_{k-a+1,k} + (1-\delta_k)t_{k-a+1,k}\}\, \chi_{b-a+1,b} & \text{if } \delta_b = 1, \\ \prod_{k=a+1}^{b-1} \{\delta_k s_{k-a+1,k} + (1-\delta_k)t_{k-a+1,k}\}\, r_{b-a+1,b} & \text{if } \delta_b = 2. \end{cases}$$

The probability of never being seen again, $\chi_{i,j}$, with $\chi_{i,n_2+1} = 1$, can be shown to be a function of $s_{i,j}$, $t_{i,j}$ and $r_{i,j}$, by first expanding $\chi_{i,j}$ to give

$$\begin{aligned} \chi_{i,j} \;=\; & (1-\phi_{i,j})(1-\lambda_{i,j}) + \phi_{i,j}(1-p_{i+1,j+1}) \cdot \\ & [(1-\phi_{i+1,j+1})(1-\lambda_{i+1,j+1}) + \phi_{i+1,j+1}(1-p_{i+2,j+2}) \cdot \\ & [(1-\phi_{i+2,j+2})(1-\lambda_{i+2,j+2}) + \phi_{i+2,j+2}(1-p_{i+3,j+3}) \cdot \\ & [\cdots[((1-\phi_{n_2,n_2})(1-\lambda_{n_2,n_2}) + \phi_{n_2,n_2}(1-p_{n_2+1,n_2+1})]\cdots]]]. \end{aligned}$$

By noting that $(1-\phi_{i,j})(1-\lambda_{i,j}) = (1-s_{i,j}-t_{i,j}-r_{i,j})$

$$\begin{aligned} \chi_{i,j} \;=\; & (1-s_{i,j}-t_{i,j}-r_{i,j}) + t_{i,j} \cdot \\ & [(1-s_{i+1,j+1}-t_{i+1,j+1}-r_{i+1,j+1}) + t_{i+1,j+1} \cdot \\ & [(1-s_{i+2,j+2}-t_{i+2,j+2}-r_{i+2,j+2}) + t_{i+2,j+2} \cdot \\ & [\cdots[(1-s_{n_2,n_2}-r_{n_2,n_2})]\cdots]]]. \end{aligned}$$

Therefore all life-histories can be parameterised in terms of **s**, **t** and **r** only.

Part two:

We derive a new exhaustive summary consisting of following terms:

- $u_j = \prod_{k=1}^{n_2-j+1} \phi_{k,k+j-1}p_{k+1,k+j}$,
  for all $j = 1,\ldots,n_2$;
- $v_{i,j} = \dfrac{(1-p_{i+1,j+1})}{p_{i+1,j+1}}$,
  for all $i = 1,\ldots,n_2-1$ and $j = i,\ldots,\min(i+n_1-1, n_2-1)$;
- $w_{i,j} = \dfrac{\chi_{i+1,j+1}}{\left(\prod_{k=1}^{n_2-j} \phi_{k+i,k+j}p_{k+i+1,k+j+1}\right)}$,
  for all $i = 1,\ldots,n_2-1$ and $j = i,\ldots,\min(i+n_1-1, n_2-1)$;

- and $x_{i,j} = \dfrac{r_{i,j}}{\left(\prod_{k=0}^{n_2-j} \phi_{k+i,k+j} p_{k+i+1,k+j+1}\right)}$,
  for all $i = 1, \ldots, n_2$ and $j = i, \ldots, \min(i + n_1 - 1, n_2)$.

We can reparameterise the original exhaustive summary consisting of the life-histories, when there are $n_1 = n_2 = 2$ years of marking and recapture/recovery, as

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(111) \\ Pr(101) \\ Pr(011) \\ Pr(110) \\ Pr(112) \\ Pr(120) \\ Pr(102) \\ Pr(012) \\ Pr(100) \\ Pr(010) \end{bmatrix} = \begin{bmatrix} \phi_{1,1}p_{2,2}\phi_{2,2}p_{3,3} \\ \phi_{1,1}\bar{p}_{2,2}\phi_{2,2}p_{3,3} \\ \phi_{1,2}p_{2,3} \\ \phi_{1,1}p_{2,2}\chi_{2,2} \\ \phi_{1,1}p_{2,2}\bar{\phi}_{2,2}\lambda_{2,2} \\ \bar{\phi}_{1,1}\lambda_{1,1} \\ \phi_{1,1}\bar{p}_{2,2}\bar{\phi}_{2,2}\lambda_{2,2} \\ \bar{\phi}_{1,2}\lambda_{1,2} \\ \chi_{1,1} \\ \bar{\phi}_{1,2}\bar{\lambda}_{1,2} + \phi_{1,2}\bar{p}_{2,3} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_1 v_{1,1} \\ u_2 \\ u_1 w_{1,1} \\ u_1 x_{2,2} \\ u_1 x_{1,1} \\ u_1 v_{1,1} x_{2,2} \\ u_2 x_{1,2} \\ k_9 \\ 1 - u_2 - u_2 x_{1,2} \end{bmatrix},$$

where $k_9 = 1 - u_1 - u_1 v_{1,1} - u_1 w_{1,1} - u_1 x_{2,2} - u_1 x_{1,1} - u_1 v_{1,1} x_{2,2}$ and $\bar{\phi} = (1 - \phi)$ and. The reparameterisation is given as $\boldsymbol{\kappa}_{uvwx} = [u_1, u_2, v_{1,1}, w_{1,1}, x_{1,1}, x_{1,2}, x_{2,2}]^T$ where the derivative matrix is

$$\left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\kappa}_{uvwx}}\right] = \begin{bmatrix} 1 & v_{1,1} & 0 & w_{1,1} & x_{2,2} & x_{1,1} & v_{1,1}x_{2,2} & 0 & D_{1,9} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & x_{1,2} & 0 & -1 - x_{1,2} \\ 0 & u_1 & 0 & 0 & 0 & 0 & u_1 x_{2,2} & 0 & -u_1(1 + x_{2,2}) & 0 \\ 0 & 0 & 0 & u_1 & 0 & 0 & 0 & 0 & -u_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & u_1 & 0 & 0 & -u_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & u_2 & 0 & -u_2 \\ 0 & 0 & 0 & 0 & u_1 & 0 & u_1 v_{1,1} & 0 & -u_1(1 + v_{1,1}) & 0 \end{bmatrix},$$

has full rank 7 where $D_{1,9} = -1 - v_{1,1} + v_{1,1}x_{2,2} - x_{2,2} - x_{1,1} - w_{1,1}$. A modified PLUR decomposition of $[\partial\boldsymbol{\kappa}/\partial\boldsymbol{\kappa}_{uvwx}]$ shows this is valid for all values of $u_1, u_2, v_{1,1}$ and $w_{1,1}$ as long as $u_1 = \phi_{1,1}p_{2,2}\phi_{2,2}p_{3,3}$ and $u_2 = \phi_{1,2}p_{2,3}$ are non-zero. This only occurs at boundary values, which have been excluded from the parameter space. Therefore by Theorem 2.3, when $n_1 = n_2 = 2$, $\boldsymbol{\kappa}_{uvwx}$ is an alternative exhaustive summary for the model. Now consider extending the model firstly by adding another year of recapture

so that $n_2 = 3$, while keeping $n_1 = 2$. The original exhaustive summary is then

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(1111) \\ Pr(1011) \\ Pr(0111) \\ Pr(1100) \\ Pr(1120) \\ Pr(1200) \\ Pr(1020) \\ Pr(0120) \\ Pr(1000) \\ Pr(0101) \\ Pr(1110) \\ Pr(1010) \\ Pr(0110) \\ Pr(1101) \\ Pr(1001) \\ Pr(0100) \\ Pr(1112) \\ Pr(1102) \\ Pr(1012) \\ Pr(1002) \\ Pr(0112) \\ Pr(0102) \end{bmatrix} = \begin{bmatrix} u_1 \\ u_1 v_{1,1} \\ u_2 \\ u_1 w_{1,1} \\ u_1 x_{2,2} \\ u_1 x_{1,1} \\ u_1 v_{1,1} x_{2,2} \\ u_2 x_{1,2} \\ k_9 \\ u_2 v_{1,2} \\ u_1 w_{2,2} \\ u_1 v_{1,1} w_{2,2} \\ u_2 w_{1,2} \\ u_1 v_{2,2} \\ u_1 v_{1,1} v_{2,2} \\ k_{16} \\ u_1 x_{3,3} \\ u_1 v_{2,2} x_{3,3} \\ u_1 v_{1,1} x_{3,3} \\ u_1 v_{1,1} v_{2,2} x_{3,3} \\ u_2 x_{2,3} \\ u_2 v_{1,2} x_{2,3} \end{bmatrix},$$

where $k_9 = 1 - u_1 - u_1 v_{1,1} - u_1 w_{1,1} - u_1 x_{2,2} - u_1 x_{1,1} - u_1 v_{1,1} x_{2,2} - u_1 w_{2,2} - u_1 v_{1,1} w_{2,2} - u_1 v_{2,2} - u_1 v_{1,1} v_{2,2} - u_1 x_{3,3} - u_1 v_{2,2} x_{3,3} - u_1 v_{1,1} x_{3,3} - u_1 v_{1,1} v_{2,2} x_{3,3}$ and $k_{16} = 1 - u_2 - u_2 x_{1,2} - u_2 v_{1,2} - u_2 w_{1,2} - u_2 x_{2,3} - u_2 v_{1,2} x_{2,3}$. We write $\boldsymbol{\kappa}_{uvwx} = [u_1, u_2, v_{1,1}, v_{1,2}, v_{2,2}, w_{1,1}, w_{1,2}, w_{2,2}, x_{1,1}, x_{1,2}, x_{2,2}, x_{2,3}, x_{3,3}]^T$. The two-stage extension theorem Theorem

3.1 is applied next. The first stage involves the exhaustive summary terms

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} u_1 \\ u_1 v_{1,1} \\ u_2 \\ u_1 w_{1,1} \\ u_1 x_{2,2} \\ u_1 x_{1,1} \\ u_1 v_{1,1} x_{2,2} \\ u_2 x_{1,2} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [u_1, u_2, v_{1,1}, w_{1,1}, x_{1,1}, x_{1,2}, x_{2,2}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_1}{\partial \boldsymbol{\theta}_1}\right] = \begin{bmatrix} 1 & v_{1,1} & 0 & w_{1,1} & x_{2,2} & x_{1,1} & v_{1,1}x_{2,2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & x_{1,2} \\ 0 & u_1 & 0 & 0 & 0 & 0 & u_1 x_{2,2} & 0 \\ 0 & 0 & 0 & u_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 & u_1 & 0 & u_1 v_{1,1} & 0 \end{bmatrix},$$

has full rank 7. The second stage examines the remaining exhaustive summary terms

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} k_9 \\ u_2 v_{1,2} \\ u_1 w_{2,2} \\ u_1 v_{1,1} w_{2,2} \\ u_2 w_{1,2} \\ u_1 v_{2,2} \\ u_1 v_{1,1} v_{2,2} \\ k_{16} \\ u_1 x_{3,3} \\ u_1 v_{2,2} x_{3,3} \\ u_1 v_{1,1} x_{3,3} \\ u_1 v_{1,1} v_{2,2} x_{3,3} \\ u_2 x_{2,3} \\ u_2 v_{1,2} x_{2,3} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_2 = [v_{1,2}, v_{2,2}, w_{1,2}, w_{2,2}, x_{2,3}, x_{3,3}]^T$. The derivative matrix $[\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$ has full rank 6. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of years of recapture. Adding a year of marking so $n_1 = 3$ while $n_2 = 3$ adds the following exhaustive summary terms

$$\boldsymbol{\kappa}_3 = \begin{bmatrix} Pr(0011) \\ Pr(0010) \\ Pr(0012) \end{bmatrix} = \begin{bmatrix} \phi_{1,3}p_{2,4} \\ 1-\phi_{1,3}p_{2,4} \\ (1-\phi_{1,3})\lambda_{1,3} \end{bmatrix} = \begin{bmatrix} u_3 \\ 1-u_3-u_3x_{1,3} \\ u_1x_{1,3} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [u_3, x_{1,3}]$. The derivative matrix

$$\left[\frac{\partial\boldsymbol{\kappa}_3}{\partial\boldsymbol{\theta}_3}\right] = \begin{bmatrix} 1 & -1-x_{1,3} & x_{1,3} \\ 0 & -u_3 & u_3 \end{bmatrix},$$

has full rank 2. Therefore $\boldsymbol{\kappa}_{uvwx}$ is an exhaustive summary for any dimension.

Part three:

This part involves checking whether the derivative matrix $[\partial\boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r})/\partial(\mathbf{s},\mathbf{t},\mathbf{r})]$ is full rank and then using the two-stage extension theorem of Theorem 3.1 to show it is always full rank for larger dimensions. Starting with $n_1 = n_2 = 2$ we can reparameterise $\boldsymbol{\kappa}_{uvwx}$ in terms of $\mathbf{s}$, $\mathbf{t}$ and $\mathbf{r}$ to get

$$\boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r}) = \begin{bmatrix} u_1 \\ v_{1,1} \\ w_{1,1} \\ u_2 \\ x_{1,1} \\ x_{1,2} \\ x_{2,2} \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2} \\ t_{1,1}/s_{1,1} \\ (1-s_{2,2}-r_{2,2})/s_{2,2} \\ s_{1,2} \\ r_{1,1}/s_{1,1}s_{2,2} \\ r_{1,2}/s_{1,2} \\ r_{2,2}/s_{2,2} \end{bmatrix}, \tag{B.2}$$

with parameter set $\boldsymbol{\kappa}_{str} = [s_{1,1}, s_{1,2}, s_{2,2}, t_{1,1}, r_{1,1}, r_{1,2}, r_{2,2}]^T$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r})}{\partial \boldsymbol{\kappa}_{str}}\right] = \begin{bmatrix} s_{2,2} & -\frac{t_{1,1}}{s_{1,1}^2} & 0 & 0 & -\frac{r_{1,1}}{s_{1,1}^2 s_{2,2}} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \frac{-r_{1,2}}{s_{1,2}} & 0 \\ s_{1,1} & 0 & \frac{-1}{s_{2,2}} & -\frac{1-s_{2,2}-r_{2,2}}{s_{2,2}^2} & 0 & -\frac{r_{1,1}}{s_{1,1}s_{2,2}^2} & 0 & -\frac{r_{2,2}}{s_{2,2}^2} \\ 0 & \frac{1}{s_{1,1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{s_{2,2}s_{1,1}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{s_{1,2}} & 0 \\ 0 & 0 & -\frac{1}{s_{2,2}} & 0 & 0 & 0 & \frac{1}{s_{2,2}} \end{bmatrix},$$

has full rank 7. A modified PLUR decomposition of $[\partial \boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r})/\partial \boldsymbol{\kappa}_{str}]$ shows this the model remains full rank for any value of $s_{1,1}, s_{1,2}, s_{2,2}, t_{1,1}, r_{1,1}, r_{1,2}$ or $r_{2,2}$. Therefore $\boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r})$ is an exhaustive summary when $n_1 = n_2 = 2$. If we extend the model to add another year of recapture, the parameter set is

$$\boldsymbol{\kappa}_{uvwx}(\mathbf{s},\mathbf{t},\mathbf{r}) = \begin{bmatrix} u_1 \\ u_2 \\ v_{1,1} \\ v_{1,2} \\ v_{2,2} \\ w_{1,1} \\ w_{1,2} \\ w_{2,2} \\ x_{1,1} \\ x_{1,2} \\ x_{2,2} \\ x_{2,3} \\ x_{3,3} \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2}s_{3,3} \\ s_{1,2}s_{2,3} \\ t_{1,1}/s_{1,1} \\ t_{1,2}/s_{1,2} \\ t_{2,2}/s_{2,2} \\ (1-s_{3,3}-s_{3,3})/s_{3,3} \\ (1-s_{2,3}-r_{2,3})/s_{2,3} \\ \{(1-s_{2,2}-t_{2,2}-r_{2,2})+t_{2,2}(1-s_{3,3})\}/s_{2,2}s_{3,3} \\ r_{1,1}/s_{1,1}s_{2,2}s_{3,3} \\ r_{1,2}/s_{1,2}s_{2,3} \\ r_{2,2}/s_{2,2}s_{3,3} \\ r_{2,3}/s_{2,3} \\ r_{3,3}/s_{3,3} \end{bmatrix},$$

with parameters $\boldsymbol{\kappa}_{str} = [s_{1,1}, s_{1,2}, s_{2,2}, s_{2,3}, s_{3,3}, t_{1,1}, t_{1,2}, t_{2,2}, r_{1,1}, r_{1,2}, r_{2,2}, r_{2,3}, r_{3,3}]^T$. Note that the terms $u_2, v_{2,2}w_{1,2}, x_{1,2}$ and $x_{2,3}$ are identical to $u_1, v_{1,1}, w_{1,1}, x_{1,1}$ and $x_{2,2}$ respectively if $s_{1,1}$ is re-labelled as $s_{1,2}$, $s_{2,2}$ as $s_{2,3}$, $t_{1,1}$ as $t_{1,2}$, $r_{1,1}$ as $r_{1,2}$, and $r_{2,2}$

as $r_{2,3}$. This can then form stage one of the two-stage extension theorem with

$$\boldsymbol{\kappa}_c = \begin{bmatrix} u_2 \\ v_{1,2} \\ w_{1,2} \\ x_{1,2} \\ x_{2,2} \end{bmatrix} = \begin{bmatrix} s_{1,2}s_{2,3} \\ t_{1,2}/s_{1,2} \\ (1-s_{2,3}-r_{2,3})/s_{2,3} \\ r_{1,2}/s_{1,2}s_{2,3} \\ r_{2,3}/s_{2,3} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [s_{1,2}, s_{2,3}, t_{1,2}, r_{1,2}, r_{2,3}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_1}{\partial \boldsymbol{\theta}_1}\right] = \begin{bmatrix} s_{2,3} & -\frac{t_{1,2}}{s_{1,2}^2} & 0 & -\frac{r_{1,2}}{s_{1,2}^2 s_{2,3}} & 0 \\ s_{1,2} & 0 & -\frac{1}{s_{2,3}} - \frac{1-s_{2,3}-r_{2,3}}{s_{2,3}^2} & -\frac{r_{1,2}}{s_{1,2}s_{2,3}^2} & \frac{r_{2,3}}{s_{2,3}^2} \\ 0 & \frac{1}{s_{1,2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{s_{1,2}s_{2,3}} & 0 \\ 0 & 0 & \frac{-1}{s_{2,3}} & 0 & \frac{1}{s_{2,3}} \end{bmatrix},$$

has full rank 5. The second stage involves the exhaustive summary terms

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} u_1 \\ v_{1,1} \\ v_{2,2} \\ w_{1,1} \\ w_{2,2} \\ x_{1,1} \\ x_{2,2} \\ x_{3,3} \end{bmatrix} = \begin{bmatrix} s_{1,1}s_{2,2}s_{3,3} \\ t_{1,1}/s_{1,1} \\ t_{2,2}/s_{2,2} \\ (1-s_{3,3})/s_{3,3} \\ \left\{(1-s_{2,2}-t_{2,2}-r_{2,2}) + t_{2,2}\,(1-s_{3,3})\right\}/s_{2,2}s_{3,3} \\ r_{1,1}/s_{1,1}s_{2,2}s_{3,3} \\ r_{2,2}/s_{2,2}s_{3,3} \\ r_{3,3}/s_{3,3} \end{bmatrix},$$

with the parameter set $\boldsymbol{\theta}_2 = [s_{1,1}s_{2,2}, s_{3,3}, t_{1,1}, t_{2,2}, r_{1,1}, r_{2,2}, r_{3,3}]^T$. The derivative matrix $[\partial \boldsymbol{\kappa}_2/\partial \boldsymbol{\theta}_2]$ has full rank 5. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of years of recapture. Adding a year of marking so $n_1 = 3$ while $n_2 = 3$ adds the following exhaustive summary terms

$$\boldsymbol{\kappa}_3 = \begin{bmatrix} u_3 \\ x_{1,3} \end{bmatrix} = \begin{bmatrix} s_{1,3} \\ r_{1,3}/s_{1,3} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [s_{1,3}, x_{1,3}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_3}{\partial \boldsymbol{\theta}_3}\right] = \begin{bmatrix} 1 & \frac{-r_{1,3}}{s_{1,3}^2} \\ 0 & \frac{1}{s_{1,3}} \end{bmatrix},$$

has full rank 2. Therefore $\mathbf{s}$, $\mathbf{t}$ and $\mathbf{r}$ form an exhaustive summary for any dimension. This then completes the proof of Theorem 4.1. □

B.3 Proof of the intrinsic fidelity model results

We prove the results of Table 4.8 in three steps:

- In part one, we show that the original exhaustive summary where $n_2 = 2$ consisting of the fidelity life-histories can be reparameterised in terms of $[\mathbf{t}, \mathbf{v}, \mathbf{u}, \mathbf{w}]$, utilising the reparameterisation theorem of Theorem 2.3. We can then show that this simpler exhaustive summary is full rank.
- In part two, we consider extending the model by an extra recapture and recovery occasion so that $n_2 = 3$. We show this extension is full rank by splitting the exhaustive summary into two parts by considering relabelling some of the life-histories. We then achieve the final result for the fully time-dependent case.
- In part three, we consider other models which are not fully time-dependent and derive results for these models.

We assume for these results that none of the parameters are on boundary values so our parameter space only includes the case where none of the parameters are equal to 0 or 1. If some of the parameters are on boundary values, this could change the parameter redundancy of the model.

Part one:

Begin with the case where $n_1 = n_2 = 2$ with two recapture occasions and three recovery occasions and where all of the parameters are time-dependent. All of the

possible life-histories are given below:

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(300) \\ Pr(130) \\ Pr(120) \\ Pr(113) \\ Pr(112) \\ Pr(111) \\ Pr(110) \\ Pr(103) \\ Pr(102) \\ Pr(101) \\ Pr(100) \\ Pr(030) \\ Pr(013) \\ Pr(012) \\ Pr(011) \\ Pr(010) \\ Pr(003) \\ Pr(001) \end{bmatrix} = \begin{bmatrix} (1-S_1)\lambda_1 \\ S_1F_1p_2(1-S_2)\lambda_2 \\ S_1\gamma_{1,2}(1-S_2)\lambda_2 \\ S_1F_1p_2S_2F_2p_3(1-S_3)\lambda_3 \\ S_1F_1p_2S_2\gamma_{2,3}(1-S_3)\lambda_3 \\ S_1F_1p_2S_2F_2p_3\chi_3 \\ S_1F_1p_2\chi_2 \\ S_1F_1(1-p_2)S_2F_2p_3(1-S_3)\lambda_3 \\ S_1S_2\gamma_{1,3}(1-S_3)\lambda_3 \\ S_1F_1(1-p_2)S_2F_2p_3\chi_3 \\ \chi_1 \\ (1-S_2)\lambda_2 \\ S_2F_2p_3(1-S_3)\lambda_3 \\ S_2\gamma_{2,3}(1-S_3)\lambda_3 \\ S_2F_2p_3\chi_3 \\ \chi_2 \\ (1-S_3)\lambda_3 \\ \chi_3 \end{bmatrix},$$

where

$$\chi_j = (1-S_j)(1-\lambda_j) + S_j(1-F_j)\Psi_{j+1} + S_jF_j(1-p_{j+1})\chi_{j+1},$$

with $\chi_{n_2+1} = (1-S_{n_2+1})(1-\lambda_{n_2+1}) + S_{n_2+1}$, where

$$\Psi_j = (1-S_j)(1-\lambda_j) + S_j\Psi_{j+1},$$

with $\Psi_{n_2+1} = 1$, and

$$\gamma_{a,b} = (1-F_a) + F_a(1-p_{a+1})\gamma_{a+1,b},$$

with $\gamma_{b,b} = 1$. These formulae were previously given in Section 4.8.

Now we can reparameterise this model by using the following parameters:

$$\begin{array}{ll} t_i = F_i p_{i+1} & \text{for } i = 1, \ldots, n_2, \\ u_i = F_i(1 - p_{i+1}) & \text{for } i = 1, \ldots, n_2 - 1, \\ v_i = (1 - S_i)\lambda_i & \text{for } i = 1, \ldots, n_2 + 1, \text{and} \\ w_i = (1 - S_i)(1 - \lambda_i) & \text{for } i = 1, \ldots, n_2. \end{array}$$

Further note that we can obtain individual original parameters as

$$\begin{aligned} t_i + u_i &= F_i p_{i+1} + F_i(1 - p_{i+1}) \\ &= F_i, \end{aligned}$$

and

$$\begin{aligned} 1 - v_i - w_i &= 1 - (1 - S_i)\lambda_i - (1 - S_i)(1 - \lambda_i) \\ &= 1 - (\lambda_i - S_i\lambda_i) - (1 - S_i - \lambda_i + S_i\lambda_i) \\ &= S_i. \end{aligned}$$

This means the exhaustive summary can alternatively be written as

$$\boldsymbol{\kappa}_s = \begin{bmatrix} Pr(300) \\ Pr(130) \\ Pr(120) \\ Pr(113) \\ Pr(112) \\ Pr(111) \\ Pr(110) \\ Pr(103) \\ Pr(102) \\ Pr(101) \\ Pr(100) \\ Pr(030) \\ Pr(013) \\ Pr(012) \\ Pr(011) \\ Pr(010) \\ Pr(003) \\ Pr(001) \end{bmatrix} = \begin{bmatrix} v_1 \\ (1 - v_1 - w_1)t_1 v_2 \\ (1 - v_1 - w_1)\gamma_{1,2} v_2 \\ (1 - v_1 - w_1)t_1(1 - v_2 - w_2)t_2 v_3 \\ (1 - v_1 - w_1)t_1(1 - v_2 - w_2)\gamma_{2,3} v_3 \\ (1 - v_1 - w_1)t_1(1 - v_2 - w_2)t_2\chi_3 \\ (1 - v_1 - w_1)t_1\chi_2 \\ (1 - v_1 - w_1)u_1(1 - v_2 - w_2)t_2 v_3 \\ (1 - v_1 - w_1)(1 - v_2 - w_2)\gamma_{1,3} v_3 \\ (1 - v_1 - w_1)u_1(1 - v_2 - w_2)t_2\chi_3 \\ \chi_1 \\ v_2 \\ (1 - v_2 - w_2)t_2 v_3 \\ (1 - v_2 - w_2)\gamma_{2,3} v_3 \\ (1 - v_2 - w_2)t_2\chi_3 \\ \chi_2 \\ v_3 \\ \chi_3 \end{bmatrix}, \tag{B.3}$$

where

$$\gamma_{i,j} = (1 - t_i - u_i) + u_i\gamma_{i+1,j}, \tag{B.4}$$

with $\gamma_{j,j} = 1$, and

$$\chi_j = w_j + (1 - v_j - w_j)(1 - t_j - u_j)\Psi_{j+1} + (1 - v_j - w_j)u_j\chi_{j+1}, \tag{B.5}$$

with $\chi_{n_2+1} = w_{n_2+1} + (1 - v_{n_2+1} - w_{n_2+1}) = 1 - v_{n_2+1}$, where

$$\Psi_j = w_j + (1 - v_j - w_j)\Psi_{j+1},$$

with $\Psi_{n_2+1} = w_{n_2+1} + (1 - v_{n_2+1} - w_{n_2+1}) = 1 - v_{n_2+1}$. Observe that we do not require a u_{n_2} term for this reparameterisation as

$$\begin{aligned}
\chi_{n_2} &= w_{n_2} + (1 - v_{n_2} - w_{n_2})(1 - t_{n_2} - u_{n_2})\Psi_{n_2+1} + (1 - v_{n_2} - w_{n_2})u_{n_2}\chi_{n_2+1} \\
&= w_{n_2} + (1 - v_{n_2} - w_{n_2})(1 - t_{n_2} - u_{n_2})(1 - v_{n_2+1}) + (1 - v_{n_2} - w_{n_2})u_{n_2}(1 - v_{n_2+1}) \\
&= w_{n_2} + (1 - v_{n_2} - w_{n_2})(1 - v_{n_2+1})\left\{(1 - t_{n_2} - u_{n_2}) + u_{n_2}\right\} \\
&= w_{n_2} + (1 - v_{n_2} - w_{n_2})(1 - v_{n_2+1})(1 - t_{n_2}).
\end{aligned}$$

This algebra means we can eventually reparameterise the original model containing 10 parameters by using 8 different parameters given as $\boldsymbol{\theta}_s = [t_1,\ t_2,\ u_1,\ v_1,\ v_2,\ v_3,\ w_1,\ w_2]^T$. This gives the derivative matrix, $\mathbf{D}_s$, when we differentiate the simpler exhaustive summary, $\boldsymbol{\kappa}_s$, with respect to its parameters as

$$\begin{bmatrix}
0 & (1 - v_1 - w_1)v_2 & -(1 - v_1 - w_1)v_2 & (1 - v_1 - w_1)(1 - v_2 - w_2)t_2v_3 & \cdots & 0 & 0 \\
0 & 0 & 0 & (1 - v_1 - w_1)t_1(1 - v_2 - w_2)v_3 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
1 & -t_1v_2 & -(1 - t_1)v_2 & -(1 - v_2 - w_2)t_2v_3 & \cdots & 0 & 0 \\
0 & (1 - v_1 - w_1)t_1 & (1 - v_1 - w_1)(1 - t_1) & -(1 - v_1 - w_1)t_1t_2v_3 & \cdots & 0 & 0 \\
0 & 0 & 0 & (1 - v_1 - w_1)t_1(1 - v_2 - w_2)t_2 & \cdots & 1 & -1 \\
0 & -t_1v_2 & -(1 - t_1)v_2 & -t_1(1 - v_2 - w_2)t_2v_3 & \cdots & 0 & 0 \\
0 & 0 & 0 & -(1 - v_1 - w_1)t_1t_2v_3 & \cdots & 0 & 0
\end{bmatrix}. \tag{B.6}$$

This derivative matrix has rank 8 and is full rank. This shows that the original exhaustive summary containing 2 recapture and 3 recovery occasions is full rank. A modified PLUR decomposition of $[\partial\boldsymbol{\kappa}_s/\partial(\mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w})]$ shows this the model remains full rank for any non-boundary values.

Part two:

Now let us add an additional recapture and recovery occasion so that $n_2 = 3$. Consider the partition where the exhaustive summary is in two parts where $\boldsymbol{\kappa} = [\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2]^T$. Let $\boldsymbol{\kappa}_1$ be the life-histories of the animals which are marked in the second capture occasion and beyond as

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} Pr(0300) \\ Pr(0130) \\ Pr(0120) \\ Pr(0113) \\ Pr(0112) \\ Pr(0111) \\ Pr(0110) \\ Pr(0103) \\ Pr(0102) \\ Pr(0101) \\ Pr(0100) \\ Pr(0030) \\ Pr(0013) \\ Pr(0012) \\ Pr(0011) \\ Pr(0010) \\ Pr(0003) \\ Pr(0001) \end{bmatrix} = \begin{bmatrix} (1-S_2)\lambda_2 \\ S_2F_2p_3(1-S_3)\lambda_3 \\ S_2\gamma_{2,3}(1-S_3)\lambda_3 \\ S_2F_2p_3S_3F_3p_4(1-S_4)\lambda_4 \\ S_2F_2p_3S_3\gamma_{3,4}(1-S_4)\lambda_4 \\ S_2F_2p_3S_3F_3p_4\chi_4 \\ S_2F_2p_3\chi_3 \\ S_2F_2(1-p_3)S_3F_3p_4(1-S_4)\lambda_4 \\ S_2S_3\gamma_{2,4}(1-S_4)\lambda_4 \\ S_2F_2(1-p_3)S_3F_3p_4\chi_4 \\ \chi_2 \\ (1-S_3)\lambda_3 \\ S_3F_3p_4(1-S_4)\lambda_4 \\ S_3\gamma_{3,4}(1-S_4)\lambda_4 \\ S_3F_3p_4\chi_4 \\ \chi_3 \\ (1-S_4)\lambda_4 \\ \chi_4 \end{bmatrix}.$$

$\boldsymbol{\kappa}_1$ can be reparameterised in terms of $[\mathbf{t}, \mathbf{v}, \mathbf{u}, \mathbf{w}]$ as

$$\boldsymbol{\kappa}_{s,1} = \begin{bmatrix} Pr(0300) \\ Pr(0130) \\ Pr(0120) \\ Pr(0113) \\ Pr(0112) \\ Pr(0111) \\ Pr(0110) \\ Pr(0103) \\ Pr(0102) \\ Pr(0101) \\ Pr(0100) \\ Pr(0030) \\ Pr(0013) \\ Pr(0012) \\ Pr(0011) \\ Pr(0010) \\ Pr(0003) \\ Pr(0001) \end{bmatrix} = \begin{bmatrix} v_2 \\ (1-v_2-w_2)t_2v_3 \\ (1-v_2-w_2)\gamma_{2,3}v_3 \\ (1-v_2-w_2)t_2(1-v_3-w_3)t_3v_4 \\ (1-v_2-w_2)t_2(1-v_3-w_3)\gamma_{3,4}v_4 \\ (1-v_2-w_2)t_2(1-v_3-w_3)t_3\chi_4 \\ (1-v_2-w_2)t_2\chi_3 \\ (1-v_2-w_2)u_2(1-v_3-w_3)t_3v_4 \\ (1-v_2-w_2)(1-v_3-w_3)\gamma_{2,4}v_4 \\ (1-v_2-w_2)u_2(1-v_3-w_3)t_3\chi_4 \\ \chi_2 \\ v_3 \\ (1-v_3-w_3)t_3v_4 \\ (1-v_3-w_3)\gamma_{3,4}v_4 \\ (1-v_3-w_3)t_3\chi_4 \\ \chi_3 \\ v_4 \\ \chi_4 \end{bmatrix}, \tag{B.7}$$

where $\gamma_{i,j}$ and χ_i are as given in Equations (B.4) and (B.5) respectively. If we compare $\boldsymbol{\kappa}_{s,1}$ above in Equation (B.7) to $\boldsymbol{\kappa}_s$ in Equation (B.3), we find that these two exhaustive summaries are the same if we use a relabelling system on $\boldsymbol{\kappa}_{s,1}$ so that

$$\begin{aligned} t_3 &\mapsto t_2 \\ t_2 &\mapsto t_1 \\ u_2 &\mapsto u_1 \\ v_4 &\mapsto v_3 \\ v_3 &\mapsto v_2 \\ v_2 &\mapsto v_1 \\ w_3 &\mapsto w_2 \\ w_2 &\mapsto w_1. \end{aligned}$$

This relabelling further means that the expressions χ and γ are given as $\chi_i \mapsto \chi_{i-1}$ and $\gamma_{i,j} \mapsto \gamma_{i-1,j-1}$. The derivative matrix $\mathbf{D}_{s,1} = [\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_{s,1}]$ with respect to the

parameters $\boldsymbol{\theta}_{s,1} = [t_1,\ t_2,\ u_1,\ v_1,\ v_2,\ v_3,\ w_1,\ w_2]^T$ is the same as given in Equation (B.6) which is shown to be full rank.

Now consider the remaining exhaustive terms in $\boldsymbol{\kappa}_2$ where

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} Pr(3000) \\ Pr(1300) \\ Pr(1200) \\ Pr(1130) \\ Pr(1120) \\ Pr(1113) \\ Pr(1112) \\ Pr(1111) \\ Pr(1110) \\ Pr(1103) \\ Pr(1102) \\ Pr(1101) \\ Pr(1100) \\ Pr(1030) \\ Pr(1020) \\ Pr(1013) \\ Pr(1012) \\ Pr(1011) \\ Pr(1010) \\ Pr(1003) \\ Pr(1002) \\ Pr(1001) \\ Pr(1000) \end{bmatrix} = \begin{bmatrix} (1-S_1)\lambda_1 \\ S_1F_1p_2(1-S_2)\lambda_2 \\ S_1\gamma_{1,2}(1-S_2)\lambda_2 \\ S_1F_1p_2S_2F_2p_3(1-S_3)\lambda_3 \\ S_1F_1p_2S_2\gamma_{2,3}(1-S_3)\lambda_3 \\ S_1F_1p_2S_2F_2p_3S_3F_3p_4(1-S_4)\lambda_4 \\ S_1F_1p_2S_2F_2p_3S_3\gamma_{3,4}(1-S_4)\lambda_4 \\ S_1F_1p_2S_2F_2p_3S_3F_3p_4\chi_4 \\ S_1F_1p_2S_2F_2p_3\chi_3 \\ S_1F_1p_2S_2F_2(1-p_3)S_3F_3p_4(1-S_4)\lambda_4 \\ S_1F_1p_2S_2S_3\gamma_{2,4}(1-S_4)\lambda_4 \\ S_1F_1p_2S_2F_2(1-p_3)S_3F_3p_4\chi_4 \\ S_1F_1p_2\chi_2 \\ S_1F_1(1-p_2)S_2F_2p_3(1-S_3)\lambda_3 \\ S_1S_2\gamma_{1,3}(1-S_3)\lambda_3 \\ S_1F_1(1-p_2)S_2F_2p_3S_3F_3p_4(1-S_4)\lambda_4 \\ S_1F_1(1-p_2)S_2F_2p_3S_3\gamma_{3,4}(1-S_4)\lambda_4 \\ S_1F_1(1-p_2)S_2F_2p_3S_3F_3p_4\chi_4 \\ S_1F_1(1-p_2)S_2F_2p_3\chi_3 \\ S_1F_1(1-p_2)S_2F_2(1-p_3)S_3F_3p_4(1-S_4)\lambda_4 \\ S_1S_2S_3\gamma_{1,4}(1-S_4)\lambda_4 \\ S_1F_1(1-p_2)S_2F_2(1-p_3)S_3F_3p_4\chi_4 \\ \chi_1 \end{bmatrix},$$

which can be reparameterised as defined previously as

$$\boldsymbol{\kappa}_{s,2} = \begin{bmatrix} Pr(3000) \\ Pr(1300) \\ Pr(1200) \\ Pr(1130) \\ Pr(1120) \\ Pr(1113) \\ Pr(1112) \\ Pr(1111) \\ Pr(1110) \\ Pr(1103) \\ Pr(1102) \\ Pr(1101) \\ Pr(1100) \\ Pr(1030) \\ Pr(1020) \\ Pr(1013) \\ Pr(1012) \\ Pr(1011) \\ Pr(1010) \\ Pr(1003) \\ Pr(1002) \\ Pr(1001) \\ Pr(1000) \end{bmatrix} = \begin{bmatrix} v_1 \\ (1-v_1-w_1)t_1v_2 \\ (1-v_1-w_1)\gamma_{1,2}v_2 \\ (1-v_1-w_1)t_1(1-v_2-w_2)t_2v_3 \\ (1-v_1-w_1)t_1(1-v_2-w_2)\gamma_{2,3}v_3 \\ (1-v_1-w_1)t_1(1-v_2-w_2)t_2(1-v_3-w_3)t_3v_4 \\ (1-v_1-w_1)t_1(1-v_2-w_2)t_2(1-v_3-w_3)\gamma_{3,4}v_4 \\ (1-v_1-w_1)t_1(1-v_2-w_2)t_2(1-v_3-w_3)t_3\chi_4 \\ (1-v_1-w_1)t_1(1-v_2-w_2)t_2\chi_3 \\ (1-v_1-w_1)t_1(1-v_2-w_2)u_2(1-v_3-w_3)t_3v_4 \\ (1-v_1-w_1)t_1(1-v_2-w_2)(1-v_3-w_3)\gamma_{2,4}v_4 \\ (1-v_1-w_1)t_1(1-v_2-w_2)u_2(1-v_3-w_3)t_3\chi_4 \\ (1-v_1-w_1)t_1\chi_2 \\ (1-v_1-w_1)u_1(1-v_2-w_2)t_2v_3 \\ (1-v_1-w_1)(1-v_2-w_2)\gamma_{1,3}v_3 \\ (1-v_1-w_1)u_1(1-v_2-w_2)t_2(1-v_3-w_3)t_3v_4 \\ (1-v_1-w_1)u_1(1-v_2-w_2)t_2(1-v_3-w_3)\gamma_{3,4}v_4 \\ (1-v_1-w_1)u_1(1-v_2-w_2)t_2(1-v_3-w_3)t_3\chi_4 \\ (1-v_1-w_1)u_1(1-v_2-w_2)t_2\chi_3 \\ (1-v_1-w_1)u_1(1-v_2-w_2)u_2(1-v_3-w_3)t_3v_4 \\ (1-v_1-w_1)(1-v_2-w_2)(1-v_3-w_3)\gamma_{1,4}v_4 \\ (1-v_1-w_1)u_1(1-v_2-w_2)u_2(1-v_3-w_3)t_3\chi_4 \\ \chi_1 \end{bmatrix}. \tag{B.8}$$

This exhaustive summary contains the extra parameters $\boldsymbol{\theta}_{s,2} = [t_1, u_1, v_1, w_1]^T$ as $\boldsymbol{\kappa}_{s,1}$ contained all the remaining parameters in $\boldsymbol{\kappa}_{s,2}$ before the parameters were relabelled. If we find the derivative matrix $\mathbf{D}_{s,2}$ when you differentiate $\boldsymbol{\kappa}_{s,2}$ with respect to its parameters $\boldsymbol{\theta}_{s,2}$, this is given as

$$\begin{bmatrix} 0 & (1-v_1-w_1)v_2 & -(1-v_1-w_1)v_2 & (1-v_1-w_1)(1-v_2-w_2)t_2v_3 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 1 & -t_1v_2 & -(1-t_1)v_2 & -t_1(1-v_2-w_2)t_2v_3 & \cdots \\ 0 & -t_1v_2 & -(1-t_1)v_2 & -t_1(1-v_2-w_2)t_2v_3 & \cdots \end{bmatrix},$$

this derivative matrix has rank 4 and is full rank. As $\mathbf{D}_{s,1}$ is full rank and the extended matrix $\mathbf{D}_{s,2}$ is also full rank, this then proves the reparameterised exhaustive

summary is full rank for any increase of n_2. As we have reparameterised the model with two fewer parameters in it, the final parameter redundancy result for the fully time-dependent $S(t)$ $F(t)$ $p(t)$ $\lambda(t)$ capture-recapture-recovery fidelity model is that the model has an intrinsic parameter deficiency of 2.

Part three:

We have shown in part two that we can create a simpler exhaustive summary using the parameters [**t**, **v**, **u**, **w**]. We can now use this simpler exhaustive summary to create general intrinsic parameter redundancy results for different models where we consider relaxing some of the model parameters so they have constant probabilities instead of time-dependent ones. The simpler exhaustive summary for when $n_2 = 2$ is given by the exhaustive summary

$$\boldsymbol{\kappa} = \begin{bmatrix} t_1 \\ t_2 \\ u_1 \\ v_1 \\ v_2 \\ v_3 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} F_1 p_2 \\ F_2 p_3 \\ F_1(1-p_2) \\ (1-S_1)\lambda_1 \\ (1-S_2)\lambda_2 \\ (1-S_3)\lambda_3 \\ (1-S_1)(1-\lambda_1) \\ (1-S_2)(1-\lambda_2) \end{bmatrix}.$$

If we find the derivative matrix with respect to the parameters $\boldsymbol{\theta} = [F_1,\ F_2,\ S_1,\ S_2,\ S_3,\ p_2,\ p_3,\ \lambda_1,\ \lambda_2,\ \lambda_3]^T$ this matrix

$$\left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} p_2 & 0 & (1-p_2) & 0 & 0 & 0 & 0 & 0 \\ 0 & p_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1 & 0 & 0 & -(1-\lambda_1) & 0 \\ 0 & 0 & 0 & 0 & -\lambda_2 & 0 & 0 & -(1-\lambda_2) \\ 0 & 0 & 0 & 0 & 0 & -\lambda_3 & 0 & 0 \\ F_1 & 0 & -F_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & F_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-S_1) & 0 & 0 & -(1-S_1) & 0 \\ 0 & 0 & 0 & 0 & (1-S_2) & 0 & 0 & -(1-S_2) \\ 0 & 0 & 0 & 0 & 0 & (1-S_3) & 0 & 0 \end{bmatrix},$$

has a rank of 8 with a parameter deficiency of 2. If we let $s_1 = F_{n_2} p_{n_2+1}$ and $s_2 = (1 - S_{n_2+1})\lambda_{n_2+1}$, then the derivative matrix with respect to the new parameter set $\boldsymbol{\theta}_{\mathbf{s}} = [F_1,\, S_1,\, S_2,\, p_2,\, \lambda_1,\, \lambda_2,\, s_1,\, s_2]^T$ is equal to

$$\left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}_{\mathbf{s}}}\right] = \begin{bmatrix} p_2 & 0 & (1-p_2) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1 & 0 & 0 & -(1-\lambda_1) & 0 \\ 0 & 0 & 0 & 0 & -\lambda_2 & 0 & 0 & -(1-\lambda_2) \\ F_1 & 0 & -F_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-S_1) & 0 & 0 & -(1-S_1) & 0 \\ 0 & 0 & 0 & 0 & (1-S_2) & 0 & 0 & -(1-S_2) \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

which has full rank 8. This shows the previous result that the fully time-dependent fidelity model has a parameter deficiency of 2.

Now consider the case where either the fidelity parameters, F_i, or the recapture parameters, p_{i+1}, are constant. In this case, the above reparameterisation where $s_1 = F_{n_2} p_{n_2+1}$ now contains one constant parameter and one new parameter, e.g. if the fidelity parameters are now constant then $s_1 = F p_{n_2+1}$. However for this example, as the fidelity parameters are now constant, the parameter F can be obtained by only the use of the parameters t_1 and u_1, as $t_1 + u_1 = F p_2 + F(1-p_2) = F$. So therefore as s_1 now only contains one new parameter p_{n_2}, we lose the parameter redundancy in the model caused by the fidelity and recapture parameters being confounded. This means if the either the fidelity parameters or the recapture parameters are constant then the intrinsic parameter redundancy of the model is reduced by one.

Similar logic to above applies if either the survival parameters, S_i, or the recovery parameters, λ_{i+1}, are constant. For example if the survival parameters are constant, $s_2 = (1 - S)\lambda_{n_2+1}$ now contains only one new parameter, and we lose the parameter redundancy in the model caused by the survival and recovery parameters being confounded. This means if the either the survival parameters or the recovery parameters are constant then parameter redundancy is reduced by one.

If the fidelity parameters and the recapture parameters are both constant, $s_1 = F \times p$ can be estimated by any of the previous parameters t_i so we also lose the

confounding caused by the fidelity and recapture parameters as well. This means if both the fidelity parameters, F_i, and the recapture parameters, p_{i+1}, are constant, then parameter redundancy is again reduced by one. This also applies if both the survival parameters, S_i, and the recovery parameters, λ_i, are constant. This then completes the proof of the results given in Table 4.8. □

B.4 Proof of the intrinsic occupancy model theorem

Theorem 5.1 part a) states that the occupancy model with at least two surveys per season is not intrinsically parameter redundant. Theorem 5.1 part b) gives that a simpler exhaustive summary for the occupancy model with multiple seasons and surveys can consist of the terms $s_{t,j} = \varphi_{t-1}^{[1,1]} p_{t,j}$ (for $t = 2, \ldots, T$ and $j = 1, \ldots, k_t$) where $s_{1,j} = \varphi_0^{[1]} p_{1,j}$ (for $j = 1, \ldots, k_t$), $r_t = \varphi_{t-1}^{[1,1]} \prod_{j=1}^{k_t} p_{t,j}$ (for $t = 2, \ldots, T$ and $j = 1, \ldots, k_t$) where $r_1 = \varphi_0^{[1]} \prod_{j=1}^{k_1} p_{1,j}$ (for $j = 1, \ldots, k_t$) and $u_t = \varphi_t^{[0,1]} / \varphi_t^{[1,1]}$ (for $t = 1, \ldots, T-1$). The proof of Theorem 5.1 is split into four parts:

- In part one, we show that the original exhaustive summary consisting of the occupancy-histories can be reparameterised in terms of $[\mathbf{s}, \mathbf{r}, \mathbf{u}]$, utilising the reparameterisation theorem of Theorem 2.3.

- In part two, we create a new exhaustive summary, denoted as $\boldsymbol{\kappa}_{vwx}$, utilising the reparameterisation theorem of Theorem 2.3. This is created so that the standard extension theorem of Theorem 2.2 can be applied in order for results to be extended to any dimension.

- In part three, we show that the reparameterisation $[\mathbf{s}, \mathbf{r}, \mathbf{u}]$ is an exhaustive summary, again utilising the reparameterisation theorem of Theorem 2.3. This shows part b) of the theorem.

- In part four, we then use the standard extension theorem of Theorem 2.2 to prove part a) of the theorem using the simpler exhaustive summary we have proved in steps one to three. Starting with an initial one-season model with two surveys, we consider extending (trivially) the number of surveys in a season and then extending the number of seasons in the model. We consider three cases, where the new season has two or more surveys, the case where the new season has one survey, and the case where the new season has no new surveys, to prove the theorem's final result.

We assume that none of the parameters are on boundary values, so that our parameter space is restricted to $0 < p_{t,j} < 1$, $0 < \varphi_0^{[1]} < 1$, $0 < \varphi_t^{[1,1]} < 1$ and $0 < \varphi_t^{[0,1]} < 1$ for all

values of t and j for this theorem to apply. Theorem 2.3 states that if the derivative matrix $[\partial\boldsymbol{\kappa}(\boldsymbol{\kappa}_{\text{new}})/\partial\boldsymbol{\kappa}_{\text{new}}]$ is full rank then $\boldsymbol{\kappa}_{\text{new}}$ is a new exhaustive summary.

Part one:

The reparameterisations used in this part are $s_{t,j} = \varphi_{t-1}^{[1,1]} p_{t,j}$, $r_t = \varphi_{t-1}^{[1,1]} \prod_{j=1}^{k_t} p_{t,j}$ (where $s_{1,j} = \varphi_0^{[1]} p_{1,j}$ and $r_1 = \varphi_0^{[1]} \prod_{j=1}^{k_1} p_{1,j}$ for the $t = 1$ case) and $u_t = \varphi_t^{[0,1]}/\varphi_t^{[1,1]}$.

- We can show $\varphi_0^{[1]}$ is a product of the reparameterised parameters as

$$\varphi_0^{[1]} = \left(\frac{\prod_{j=1}^{k_1} s_{1,j}}{r_1} \right)^{1/(k_1-1)}.$$

  For example, if there were 3 surveys in the first season, we could achieve $\varphi_0^{[1]}$ by setting

$$\begin{aligned}
\left(\frac{\prod_{j=1}^{k_1} s_{1,j}}{r_1} \right)^{1/(k_1-1)} &= \left(\frac{s_{1,1} s_{1,2} s_{1,3}}{r_1} \right)^{1/(3-1)} \\
&= \left(\frac{\varphi_0^{[1]} p_{1,1} \varphi_0^{[1]} p_{1,2} \varphi_0^{[1]} p_{1,3}}{\varphi_0^{[1]} p_{1,1} p_{1,2} p_{1,3}} \right)^{1/2} \\
&= \left(\varphi_0^{[1]} \varphi_0^{[1]} \right)^{1/2} \\
&= \varphi_0^{[1]}.
\end{aligned}$$

- In a similar method to the previous bullet point, we can see that

$$\varphi_t^{[1,1]} = \left(\frac{\prod_{j=1}^{k_{t+1}} s_{t+1,j}}{r_{t+1}} \right)^{1/(k_{t+1}-1)},$$

  for any $t = 1, \ldots, T-1$.

- The $\varphi_t^{[0,1]}$ occurs in the occupancy-history when a species becomes present from a period when the species has been absent (i.e. colonisation). This parameter on its own is the only time the reparameterisation parameter $u_t = \varphi_t^{[0,1]}/\varphi_t^{[1,1]}$ is

needed in the model, i.e.

$$\begin{aligned}\varphi_t^{[0,1]} &= u_t \cdot \left(\frac{\prod_{j=1}^{k_{t+1}} s_{t+1,j}}{r_{t+1}}\right)^{1/(k_{t+1}-1)} \\ &= \left(\frac{\varphi_t^{[0,1]}}{\varphi_t^{[1,1]}}\right) \cdot \varphi_t^{[1,1]} \quad \text{(as seen in the previous bullet)} \\ &= \varphi_t^{[0,1]},\end{aligned}$$

for any $t = 1, \ldots, T-1$.

- A single $p_{t,j}$ can be found by performing the calculation

$$p_{t,j} = s_{t,j}\left(\frac{\prod_{i=1}^{k_t} s_{t,i}}{r_t}\right)^{-1/(k_t-1)}.$$

For example, if we wanted the parameter $p_{2,3}$ where there are 4 surveys in the second season, then

$$\begin{aligned}s_{t,j} \cdot \left(\frac{\prod_{i=1}^{k_t} s_{t,i}}{r_t}\right)^{-1/k_t-1} &= s_{2,3} \cdot \left(\frac{s_{2,1}s_{2,2}s_{2,3}s_{2,4}}{r_2}\right)^{-1/(4-1)} \\ &= \varphi_1^{[1,1]} p_{2,3} \cdot \left(\frac{\varphi_1^{[1,1]} p_{2,1} \varphi_1^{[1,1]} p_{2,2} \varphi_1^{[1,1]} p_{2,3} \varphi_1^{[1,1]} p_{2,4}}{\varphi_1^{[1,1]} p_{2,1} p_{2,2} p_{2,3} p_{2,4}}\right)^{-1/3} \\ &= \varphi_1^{[1,1]} p_{2,3} \cdot \left(\varphi_1^{[1,1]} \varphi_1^{[1,1]} \varphi_1^{[1,1]}\right)^{-1/3} \\ &= \varphi_1^{[1,1]} p_{2,3} \cdot \left(\frac{1}{\varphi_1^{[1,1]}}\right) \\ &= p_{2,3}.\end{aligned}$$

As every original parameter can be expressed as a product of the reparameterised parameters $[\mathbf{s}, \mathbf{r}, \mathbf{u}]$, then every original exhaustive summary term must be a product of the reparameterised parameters $[\mathbf{s}, \mathbf{r}, \mathbf{u}]$ as they only contain the original parameters.

Part two:

We derive a new exhaustive summary which consists of the following terms:

- $v = \varphi_0^{[1]} \left(\prod_{t=2}^{T} \varphi_{t-1}^{[1,1]}\right) \left(\prod_{t=1}^{T} \prod_{j=1}^{k_t} p_{t,j}\right)$,
  which represents the history of being detected during every survey in the study;

- $w_{t,j} = \dfrac{(1-p_{t,j})}{p_{t,j}}$,
  for all $j = 1, \ldots, k_t$ and $t = 1, \ldots, T$;

- and $x_{[a,b]} = \dfrac{\varphi_{a-1}^{[1,1]} \prod_{j=1}^{k_a}(1-p_{a,j}) f_{a+1} + (1-\varphi_{a-1}^{[1,1]}) g_{a+1}}{\prod_{t=a}^{b} \varphi_{t-1}^{[1,1]} \prod_{j=1}^{k_t} p_{t,j}}$,
  which is used for multiple seasons of non-detection going from season a to season b (where $2 \leq a, b \leq T$ with $a \leq b$), with the recursive functions f and g given by
  $f_t = \varphi_{t-1}^{[1,1]} \prod_{j=1}^{k_t}(1-p_{t,j}) f_{t+1} + (1-\varphi_{t-1}^{[1,1]}) g_{t+1}$, and
  $g_t = \varphi_{t-1}^{[0,1]} \prod_{j=1}^{k_t}(1-p_{t,j}) f_{t+1} + (1-\varphi_{t-1}^{[0,1]}) g_{t+1}$,
  where $f_{b+1} = 1$, and $g_{b+1} = \varphi_b^{[0,1]}/\varphi_b^{[1,1]}$ if $b < T$ with $g_{b+1} = 1$ when $b = T$.
  For $x_{[1,b]}$, then the function f is equal to
  $f_1 = \varphi_0^{[1]} \prod_{j=1}^{k_1}(1-p_{1,j}) f_2 + (1-\varphi_0^{[1]}) g_2$.
  Also observe that a simpler form arises when $a = b$, i.e. when there is only one season of non-detection between two seasons of detection, as
  $x_{[a,a]} = \dfrac{\varphi_{a-1}^{[1,1]} \prod_{j=1}^{k_a}(1-p_{a,j}) + (1-\varphi_{a-1}^{[1,1]}) \frac{\varphi_a^{[0,1]}}{\varphi_a^{[1,1]}}}{\varphi_{a-1}^{[1,1]} \prod_{j=1}^{k_a} p_{t,j}}$,
  where $\varphi_a^{[0,1]}/\varphi_a^{[1,1]}$ does not appear in $x_{[a,a]}$ when $a = T$. This complex form for $x_{[a,b]}$ stems from the fact that even though a species is not detected in the seasons $[a, b]$, the species could potentially go through periods of extinction and colonisation unknown to us due to the lack of detections in these seasons.

We can now reparameterise the original exhaustive summary consisting of the occupancy-histories, when there are two seasons of detection with two surveys per

season, as

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(11|11) \\ Pr(11|10) \\ Pr(11|01) \\ Pr(11|00) \\ Pr(10|11) \\ Pr(10|10) \\ Pr(10|01) \\ Pr(10|00) \\ Pr(01|11) \\ Pr(01|10) \\ Pr(01|01) \\ Pr(01|00) \\ Pr(00|11) \\ Pr(00|10) \\ Pr(00|01) \\ Pr(00|00) \end{bmatrix} = \begin{bmatrix} \varphi_0^{[1]} p_{1,1} p_{1,2} \varphi_1^{[1,1]} p_{2,1} p_{2,2} \\ \varphi_0^{[1]} p_{1,1} p_{1,2} \varphi_1^{[1,1]} p_{2,1} \bar{p}_{2,2} \\ \varphi_0^{[1]} p_{1,1} p_{1,2} \varphi_1^{[1,1]} \bar{p}_{2,1} p_{2,2} \\ \varphi_0^{[1]} p_{1,1} p_{1,2} (\varphi_1^{[1,1]} \bar{p}_{2,1} \bar{p}_{2,2} + \bar{\varphi}_1^{[1,1]}) \\ \varphi_0^{[1]} p_{1,1} \bar{p}_{1,2} \varphi_1^{[1,1]} p_{2,1} p_{2,2} \\ \varphi_0^{[1]} p_{1,1} \bar{p}_{1,2} \varphi_1^{[1,1]} p_{2,1} \bar{p}_{2,2} \\ \varphi_0^{[1]} p_{1,1} \bar{p}_{1,2} \varphi_1^{[1,1]} \bar{p}_{2,1} p_{2,2} \\ \varphi_0^{[1]} p_{1,1} \bar{p}_{1,2} (\varphi_1^{[1,1]} \bar{p}_{2,1} \bar{p}_{2,2} + \bar{\varphi}_1^{[1,1]}) \\ \varphi_0^{[1]} \bar{p}_{1,1} p_{1,2} \varphi_1^{[1,1]} p_{2,1} p_{2,2} \\ \varphi_0^{[1]} \bar{p}_{1,1} p_{1,2} \varphi_1^{[1,1]} p_{2,1} \bar{p}_{2,2} \\ \varphi_0^{[1]} \bar{p}_{1,1} p_{1,2} \varphi_1^{[1,1]} \bar{p}_{2,1} p_{2,2} \\ \varphi_0^{[1]} \bar{p}_{1,1} p_{1,2} (\varphi_1^{[1,1]} \bar{p}_{2,1} \bar{p}_{2,2} + \bar{\varphi}_1^{[1,1]}) \\ (\varphi_0^{[1]} \bar{p}_{1,1} \bar{p}_{1,2} + \bar{\varphi}_0^{[1]} \varphi_1^{[0,1]} / \varphi_1^{[1,1]}) \varphi_1^{[1,1]} p_{2,1} p_{2,2} \\ (\varphi_0^{[1]} \bar{p}_{1,1} \bar{p}_{1,2} + \bar{\varphi}_0^{[1]} \varphi_1^{[0,1]} / \varphi_1^{[1,1]}) \varphi_1^{[1,1]} p_{2,1} \bar{p}_{2,2} \\ (\varphi_0^{[1]} \bar{p}_{1,1} \bar{p}_{1,2} + \bar{\varphi}_0^{[1]} \varphi_1^{[0,1]} / \varphi_1^{[1,1]}) \varphi_1^{[1,1]} \bar{p}_{2,1} p_{2,2} \\ \{\varphi_0^{[1]} \bar{p}_{1,1} \bar{p}_{1,2} (\varphi_1^{[1,1]} \bar{p}_{2,1} \bar{p}_{2,2} + \bar{\varphi}_1^{[1,1]})\} + \bar{\varphi}_0^{[1]} (\varphi_1^{[0,1]} \bar{p}_{2,1} \bar{p}_{2,2} + \bar{\varphi}_1^{[0,1]}) \end{bmatrix},$$

where $\bar{\varphi} = (1 - \varphi)$, which can be expressed as

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(11|11) \\ Pr(11|10) \\ Pr(11|01) \\ Pr(11|00) \\ Pr(10|11) \\ Pr(10|10) \\ Pr(10|01) \\ Pr(10|00) \\ Pr(01|11) \\ Pr(01|10) \\ Pr(01|01) \\ Pr(01|00) \\ Pr(00|11) \\ Pr(00|10) \\ Pr(00|01) \\ Pr(00|00) \end{bmatrix} = \begin{bmatrix} v \\ vw_{2,2} \\ vw_{2,1} \\ vx_{[2,2]} \\ vw_{1,2} \\ vw_{1,2}w_{2,2} \\ vw_{1,2}w_{2,1} \\ vw_{1,2}x_{[2,2]} \\ vw_{1,1} \\ vw_{1,1}w_{2,2} \\ vw_{1,1}w_{2,1} \\ vw_{1,1}x_{[2,2]} \\ vx_{[1,1]} \\ vx_{[1,1]}w_{2,2} \\ vx_{[1,1]}w_{2,1} \\ vx_{[1,2]} \end{bmatrix}.$$

The reparameterisation is given as $\boldsymbol{\kappa}_{vwx} = [v, w_{1,1}, w_{1,2}, w_{2,1}, w_{2,2}, x_{[1,1]}, x_{[1,2]}, x_{[2,2]}]^T$ where the derivative matrix $[\partial\boldsymbol{\kappa}/\partial\boldsymbol{\kappa}_{vwx}]$ is equal to

$$\begin{bmatrix}
1 & w_{2,2} & w_{2,1} & x_{2,2} & w_{1,2} & w_{1,2}w_{2,2} & w_{1,2}w_{2,1} & w_{1,2}x_{[2,2]} & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & v & vw_{2,2} & vw_{2,1} & vx_{2,2} & \cdots \\
0 & 0 & v & 0 & 0 & 0 & vw_{1,2} & 0 & \cdots \\
0 & v & 0 & 0 & 0 & vw_{1,2} & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & v & 0 & 0 & 0 & vw_{1,2} & \cdots
\end{bmatrix}$$

$$\begin{bmatrix}
\cdots & w_{1,1} & w_{1,1}w_{2,2} & w_{1,1}w_{2,1} & w_{1,1}x_{[2,2]} & x_{[1,1]} & w_{2,2}x_{[1,1]} & w_{2,1}x_{[1,1]} & x_{[1,2]} \\
\cdots & v & vw_{2,2} & vw_{2,1} & vx_{[2,2]} & 0 & 0 & 0 & 0 \\
\cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\cdots & 0 & 0 & vw_{1,1} & 0 & 0 & 0 & vx_{[1,1]} & 0 \\
\cdots & 0 & vw_{1,1} & 0 & 0 & 0 & vx_{[1,1]} & 0 & 0 \\
\cdots & 0 & 0 & 0 & 0 & v & vw_{2,2} & vw_{2,1} & 0 \\
\cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\cdots & 0 & 0 & 0 & vw_{1,1} & 0 & 0 & 0 & 0
\end{bmatrix},$$

and it has full rank 8. A modified PLUR decomposition of $[\partial\boldsymbol{\kappa}/\partial\boldsymbol{\kappa}_{vwx}]$ shows that this is valid for all values of $u_1, u_2, v_{1,1}$ and $w_{1,1}$ apart from any of the parameters being equal to zero. This only occurs at a boundary and the parameter space has already been restricted to exclude such boundary values. Therefore by Theorem 2.3, when there are two seasons in the study with two surveys per season, $\boldsymbol{\kappa}_{vwx}$ is an alternative exhaustive summary for the model.

Now consider extending the model firstly by adding a survey for each season. The original exhaustive summary is then

$$\boldsymbol{\kappa} = \begin{bmatrix} Pr(111|111) \\ Pr(111|110) \\ Pr(111|101) \\ Pr(111|011) \\ Pr(111|100) \\ Pr(111|010) \\ Pr(111|001) \\ Pr(111|000) \\ \vdots \\ Pr(001|000) \\ Pr(000|111) \\ Pr(000|110) \\ Pr(000|101) \\ Pr(000|011) \\ Pr(000|100) \\ Pr(000|010) \\ Pr(000|001) \\ Pr(000|000) \end{bmatrix} = \begin{bmatrix} v \\ vw_{2,3} \\ vw_{2,2} \\ vw_{2,1} \\ vw_{2,2}w_{2,3} \\ vw_{2,1}w_{2,3} \\ vw_{2,1}w_{2,2} \\ vx_{[2,2]} \\ \vdots \\ vw_{1,1}w_{1,2}x_{[2,2]} \\ vx_{[1,1]} \\ vx_{[1,1]}w_{2,3} \\ vx_{[1,1]}w_{2,2} \\ vx_{[1,1]}w_{2,1} \\ vx_{[1,1]}w_{2,2}w_{2,3} \\ vx_{[1,1]}w_{2,1}w_{2,3} \\ vx_{[1,1]}w_{2,1}w_{2,2} \\ vx_{[1,2]} \end{bmatrix}.$$

This uses the reparameterisation $\boldsymbol{\kappa}_{vwx} = [v,\ w_{1,1},\ w_{1,2},\ w_{1,3},\ w_{2,1},\ w_{2,2},\ w_{2,3},\ x_{[1,1]},\ x_{[1,2]},\ x_{[2,2]}]^T$. We now use the two-stage extension theorem of Theorem 3.1. The first stage involves the exhaustive summary terms

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} v \\ vw_{1,1} \\ vw_{1,2} \\ vw_{2,1} \\ vw_{2,2} \\ vx_{[2,2]} \\ vw_{1,1}x_{[2,2]} \\ vw_{1,2}x_{[2,2]} \\ vw_{2,1}x_{[1,1]} \\ vw_{2,2}x_{[1,1]} \\ vx_{[1,2]} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [v, w_{1,1}, w_{1,2}, w_{2,1}, w_{2,2}, x_{[1,1]}, x_{[1,2]}, x_{[2,2]}]$. The derivative matrix $[\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_1]$ is equal to

$$\begin{bmatrix}
1 & w_{1,1} & w_{1,2} & w_{2,1} & w_{2,2} & x_{[2,2]} & w_{1,1}x_{[2,2]} & w_{1,2}x_{[2,2]} & w_{2,1}x_{[1,1]} & w_{2,2}x_{[1,1]} & x_{[1,2]} \\
0 & v & 0 & 0 & 0 & 0 & vx_{[2,2]} & 0 & 0 & 0 & 0 \\
0 & 0 & v & 0 & 0 & 0 & 0 & vx_{[2,2]} & 0 & 0 & 0 \\
0 & 0 & 0 & v & 0 & 0 & 0 & 0 & vx_{[1,1]} & 0 & 0 \\
0 & 0 & 0 & 0 & v & 0 & 0 & 0 & 0 & vx_{[1,1]} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & v & vw_{2,1} & vw_{2,2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v \\
0 & 0 & 0 & 0 & 0 & 0 & vw_{1,1} & vw_{1,2} & 0 & 0 & 0
\end{bmatrix}$$

and has full rank 8. The second stage examines the remaining exhaustive summary terms denoted as $\boldsymbol{\kappa}_2$ (not given here to save space), involving the terms which involve the additional parameters $\boldsymbol{\theta}_2 = [w_{1,3}, w_{2,3}]^T$. The derivative matrix $[\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$ has full rank 2, so the model can be extended in terms of additional surveys.

Next consider the extension of an extra season so there are three seasons with two surveys in each season. The original exhaustive summary is then

$$\boldsymbol{\kappa} = \begin{bmatrix}
Pr(11|11|11) \\
Pr(11|11|10) \\
Pr(11|11|01) \\
Pr(11|11|00) \\
Pr(11|10|11) \\
Pr(11|01|11) \\
Pr(11|00|11) \\
Pr(10|11|11) \\
Pr(01|11|11) \\
Pr(00|11|11) \\
\vdots \\
Pr(00|10|00) \\
Pr(11|00|00) \\
Pr(10|00|00) \\
Pr(01|00|00) \\
Pr(00|00|00)
\end{bmatrix} = \begin{bmatrix}
v \\
vw_{3,2} \\
vw_{3,1} \\
vx_{[3,3]} \\
vw_{2,2} \\
vw_{2,1} \\
vx_{[2,2]} \\
vw_{1,2} \\
vw_{1,1} \\
vx_{[1,1]} \\
\vdots \\
vx_{[1,1]}w_{2,2}x_{[3,3]} \\
vx_{[2,3]} \\
vw_{1,2}x_{[2,3]} \\
vw_{1,1}x_{[2,3]} \\
vx_{[1,3]}
\end{bmatrix}.$$

This uses the reparameterisation $\boldsymbol{\kappa}_{vwx} = [v,\ w_{1,1},\ w_{1,2},\ w_{2,1},\ w_{2,2},\ w_{3,1},\ w_{3,2},\ x_{[1,1]},\ x_{[1,2]},\ x_{[1,3]},\ x_{[2,2]},\ x_{[2,3]},\ x_{[3,3]}]^T$. We now use the two-stage extension theorem of Theorem 3.1. The first stage involves the exhaustive summary terms

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} v \\ vw_{1,1} \\ vw_{1,2} \\ vw_{2,1} \\ vw_{2,2} \\ vx_{[2,2]} \\ vw_{1,1}x_{[2,2]} \\ vw_{1,2}x_{[2,2]} \\ vw_{2,1}x_{[1,1]} \\ vw_{2,2}x_{[1,1]} \\ vx_{[1,2]} \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [v, w_{1,1}, w_{1,2}, w_{2,1}, w_{2,2}, x_{[1,1]}, x_{[1,2]}, x_{[2,2]}]$. This is the same as $\boldsymbol{\kappa}_1$ when considering the extension of more surveys, and the derivative matrix there was shown to be full rank 8. The second stage examines the remaining exhaustive summary terms denoted as $\boldsymbol{\kappa}_2$ (not given here to save space), involving the terms which involve the additional parameters $\boldsymbol{\theta}_2 = [w_{3,1}, w_{3,2}, x_{[1,3]}, x_{[2,3]}, x_{[3,3]}]^T$. The derivative matrix $[\partial\boldsymbol{\kappa}_2/\partial\boldsymbol{\theta}_2]$ has full rank 5, so the model can be extended in terms of additional seasons. Therefore, $\boldsymbol{\kappa}_{vwx}$ is an exhaustive summary for any dimension.

Part three:

This part involves checking whether the derivative matrix $[\partial\boldsymbol{\kappa}_{vwx}(\mathbf{s},\mathbf{r},\mathbf{u})/\partial(\mathbf{s},\mathbf{r},\mathbf{u})]$ is full rank and then using the two-stage theorem of Theorem 3.1 to show it is always full rank for larger dimensions. It can be shown that we can reparameterise $\boldsymbol{\kappa}_{vwx}$ in terms of $\mathbf{s}$, $\mathbf{r}$ and $\mathbf{u}$ as

- $v = \prod_{t=1}^{T} r_t$;
- $w_{t,j} = \dfrac{1}{s_{t,j}}\left\{\left(\dfrac{\prod_{m=1}^{k_t} s_{t,m}}{r_t}\right)^{1/(k_t-1)} - s_{t,j}\right\}$, for all $t = 1,\ldots,T$ and $j = 1,\ldots,k_t$; and

- $x_{[a,b]} = \dfrac{\prod_{j=1}^{k_a} \frac{r_a}{s_{a,j}} \left\{ \left(\frac{\prod_{m=1}^{k_a} s_{a,m}}{r_a} \right)^{1/(k_a-1)} - s_{a,j} \right\} f_{a+1} + \left\{ 1 - \left(\frac{\prod_{m=1}^{k_a} s_{a,m}}{r_a} \right)^{1/(k_a-1)} \right\} g_{a+1}}{\prod_{t=a}^{b} r_t},$

  for $2 \leq a, b \leq T$ with $a \leq b$, where the recursive functions f and g are equal to

  $$f_t = \prod_{j=1}^{k_t} \frac{r_t}{s_{t,j}} \left\{ \left(\frac{\prod_{m=1}^{k_t} s_{t,m}}{r_t} \right)^{1/(k_t-1)} - s_{t,j} \right\} f_{a+1} + \left\{ 1 - \left(\frac{\prod_{m=1}^{k_t} s_{t,m}}{r_t} \right)^{1/(k_t-1)} \right\} g_{a+1},$$

  and,

  $$g_t = \prod_{j=1}^{k_t} \frac{r_t u_{t-1}}{s_{t,j}} \left\{ \left(\frac{\prod_{m=1}^{k_t} s_{t,m}}{r_t} \right)^{1/(k_t-1)} - s_{t,j} \right\} f_{t+1} + (1-u_{t-1}) \left\{ 1 - \left(\frac{\prod_{m=1}^{k_t} s_{t,m}}{r_t} \right)^{1/(k_t-1)} \right\} g_{t+1},$$

  where $f_{b+1} = 1$, and $g_{b+1} = u_b$ if $b < T$ and $g_{b+1} = 1$ when $b = T$. Also observe that a simpler form arises when $a = b$, i.e. when there is only one season of non-detection between two seasons of detection, as

  $$x_{[a,a]} = \frac{\prod_{j=1}^{k_a} \frac{r_a}{s_{a,j}} \left\{ \left(\frac{\prod_{m=1}^{k_a} s_{a,m}}{r_a} \right)^{1/(k_a-1)} - s_{a,j} \right\} + \left\{ 1 - \left(\frac{\prod_{m=1}^{k_a} s_{a,m}}{r_a} \right)^{1/(k_a-1)} \right\} u_a}{r_a},$$

  where u_a does not appear in $x_{[a,a]}$ when $a = T$.

Starting with the basic two seasons with two surveys per season we can then reparameterise $\boldsymbol{\kappa}_{vwx}$ in terms of $\mathbf{s}$, $\mathbf{r}$ and $\mathbf{u}$ to get

$$\boldsymbol{\kappa}_{vwx}(\mathbf{s},\mathbf{r},\mathbf{u}) = \begin{bmatrix} v \\ w_{1,1} \\ w_{1,2} \\ w_{2,1} \\ w_{2,2} \\ x_{[1,1]} \\ x_{[1,2]} \\ x_{[2,2]} \end{bmatrix} = \begin{bmatrix} r_1 r_2 \\ \left(s_{1,1}^2 s_{1,2}\right)/r_1 - 1 \\ \left(s_{1,1} s_{1,2}^2\right)/r_1 - 1 \\ \left(s_{2,1}^2 s_{2,2}\right)/r_2 - 1 \\ \left(s_{2,1} s_{2,2}^2\right)/r_2 - 1 \\ \left(\frac{s_{1,1}^2 s_{1,2}}{r_1} - 1\right)\left(\frac{s_{1,1} s_{1,2}^2}{r_1} - 1\right) + \frac{1}{r_1}\left(1 - \frac{s_{1,1} s_{1,2}}{r_1}\right) u_1 \\ c_7 \\ \left(\frac{s_{2,1}^2 s_{2,2}}{r_2} - 1\right)\left(\frac{s_{2,1} s_{2,2}^2}{r_2} - 1\right) + \frac{1}{r_2}\left(1 - \frac{s_{2,1} s_{2,2}}{r_2}\right) \end{bmatrix},$$

where $c_7 = \left(\frac{s_{1,1}^2 s_{1,2}}{r_1} - 1\right)\left(\frac{s_{1,1} s_{1,2}^2}{r_1} - 1\right)\left\{\left(\frac{s_{2,1}^2 s_{2,2}}{r_2} - 1\right)\left(\frac{s_{2,1} s_{2,2}^2}{r_2} - 1\right) + \frac{1}{r_2}\left(1 - \frac{s_{2,1} s_{2,2}}{r_2}\right)\right\} + \frac{1}{r_1}\left(1 - \frac{s_{1,1} s_{1,2}}{r_1}\right)\left\{\left(u_1 \frac{s_{2,1}^2 s_{2,2}}{r_2} - 1\right)\left(\frac{s_{2,1} s_{2,2}^2}{r_2} - 1\right) + \frac{(1-u_1)}{r_2}\left(1 - \frac{s_{2,1} s_{2,2}}{r_2}\right)\right\}$, with parameter set $\boldsymbol{\theta}_{sru} = [s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}, r_1, r_2, u_1]^T$. The derivative matrix $[\partial \boldsymbol{\kappa}_{vwx}(\mathbf{s},\mathbf{r},\mathbf{u})/\partial \boldsymbol{\theta}_{sru}]$ has full rank 7. A modified PLUR decomposition of $[\partial \boldsymbol{\kappa}_{vwx}(\mathbf{s},\mathbf{r},\mathbf{u})/\partial \boldsymbol{\theta}_{sru}]$

shows that the model remains full rank apart from when some of the parameters are on boundary values, where these possibilities are excluded at the beginning of the proof. If we now extend the model to add another survey per season, the exhaustive summary becomes what is given below. To suppress notation in this exhaustive summary, we will denote

$$l_{t,j,k_t} = \frac{1}{s_{t,j}} \left\{ \left(\frac{\prod_{m=1}^{k_t} s_{t,m}}{r_t} \right)^{1/(k_t-1)} - s_{t,j} \right\},$$

where $k_t = 3$ in this addition single survey extension, and

$$m_{t,k_t} = 1 - \left(\frac{\prod_{j=1}^{k_t} s_{t,j}}{r_t} \right)^{1/(k_t-1)},$$

for $t = 1, \ldots, T$ and $j = 1, \ldots, k_t$, so that the exhaustive summary $\boldsymbol{\kappa}_{vwx}(\mathbf{s}, \mathbf{r}, \mathbf{u})$ is equal to

$$\begin{bmatrix} v \\ w_{1,1} \\ w_{1,2} \\ w_{1,3} \\ w_{2,1} \\ w_{2,2} \\ w_{1,3} \\ x_{[1,1]} \\ x_{[1,2]} \\ x_{[2,2]} \end{bmatrix} = \begin{bmatrix} r_1 r_2 \\ l_{1,1,2} \\ l_{1,2,2} \\ l_{1,3,2} \\ l_{2,1,2} \\ l_{2,2,2} \\ l_{2,3,2} \\ l_{1,1,2} l_{1,2,2} l_{1,3,2} + (m_{1,2} u_1 / r_1) \\ r_1 l_{1,1,2} l_{1,2,2} l_{1,3,2} \{ r_2 l_{2,1,2} l_{2,2,2} l_{2,3,2} + \\ (m_{2,2}/r_2) \} + m_{1,2} \{ r_2 u_1 l_{2,1,2} l_{2,2,2} l_{2,3,2} + ((1-u_1) m_{2,2}/r_2) \} \\ l_{2,1,2} l_{2,2,2} l_{2,3,2} + (m_{2,2}/r_1) \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_{sru} = [s_{1,1}, s_{1,2}, s_{1,3}, s_{2,1}, s_{2,2}, s_{2,3}, r_1, r_2, u_1]^T$. We now use the two-stage extension theorem of Theorem 3.1. Note that for the two surveys per season model that

$$\begin{aligned} l_{t,j,2} &= \frac{1}{s_{t,j}} \left\{ \frac{s_{t,1} s_{t,2}}{r_t} - s_{t,j} \right\}, \\ &= \left(s_{t,1} s_{t,2} s_{t,j} \right) / r_t - 1, \\ m_{t,2} &= 1 - \frac{s_{t,1} s_{t,2}}{r_t}, \end{aligned}$$

and for the three surveys per season model that

$$\begin{aligned} l_{t,j,3} &= \frac{1}{s_{t,j}}\left\{\sqrt{\frac{s_{t,1}s_{t,2}s_{t,3}}{r_t}} - s_{t,j}\right\}, \\ m_{t,3} &= 1 - \sqrt{\frac{s_{t,1}s_{t,2}s_{t,3}}{r_t}}. \end{aligned}$$

If we relabel $l_{t,j,2}$ as $l_{t,j,3}$ and $m_{t,2}$ as $m_{t,3}$ this can then form the first stage of the two-stage extension theorem

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} v \\ w_{1,1} \\ w_{1,2} \\ w_{1,3} \\ w_{2,1} \\ w_{2,2} \\ w_{2,3} \\ x_{[2,2]} \end{bmatrix} = \begin{bmatrix} r_1 r_2 \\ l_{1,1,3} \\ l_{1,2,3} \\ l_{1,3,3} \\ l_{2,1,3} \\ l_{2,2,3} \\ l_{2,3,3} \\ l_{2,1,3}l_{2,2,3}l_{2,3,3} + (m_{2,3}/r_2) \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [s_{1,1}, s_{1,2}, s_{1,3}, s_{2,1}, s_{2,2}, s_{2,3}, r_1, r_2]$. The derivative matrix $[\partial\boldsymbol{\kappa}_1/\partial\boldsymbol{\theta}_1]$ has full rank 8. The second stage involves the rest of the exhaustive terms as $\boldsymbol{\kappa}_2 = [x_{[1,1]}, x_{[1,2]}]^T$, but this only has the additional parameter u_1, so $\boldsymbol{\kappa}_2$ is trivially full rank. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of surveys per season.

We now extend the original model two-survey two-season model by adding another season, which has two surveys in the new season. The exhaustive summary $\boldsymbol{\kappa}_{vwx}(\mathbf{s}, \mathbf{r}, \mathbf{u})$

for this extension is equal to

$$\begin{bmatrix} v \\ w_{1,1} \\ w_{1,2} \\ w_{2,1} \\ w_{2,2} \\ w_{3,1} \\ w_{3,2} \\ x_{[1,1]} \\ x_{[1,2]} \\ x_{[1,3]} \\ x_{[2,2]} \\ x_{[2,3]} \\ x_{[3,3]} \end{bmatrix} = \begin{bmatrix} r_1 r_2 r_3 \\ \left(s_{1,1}^2 s_{1,2}\right)/r_1 - 1 \\ \left(s_{1,1} s_{1,2}^2\right)/r_1 - 1 \\ \left(s_{2,1}^2 s_{1,2}\right)/r_2 - 1 \\ \left(s_{2,1} s_{1,2}^2\right)/r_2 - 1 \\ \left(s_{3,1}^2 s_{1,2}\right)/r_3 - 1 \\ \left(s_{3,1} s_{1,2}^2\right)/r_3 - 1 \\ l_{1,1,2} l_{1,2,2} + (m_{1,2} u_1/r_1) \\ r_1 l_{1,1,2} l_{1,2,2} \{r_2 l_{2,1,2} l_{2,2,2} + \\ (m_{2,2}/r_2)\} + m_{1,2}/r_1 \{u_1 r_2 l_{2,1,2} l_{2,2,2} + ((1-u_1) m_{2,2} u_2/r_2)\} \\ c_{10} \\ l_{2,1,2} l_{2,2,2} + (m_{2,2} u_2/r_2) \\ r_2 l_{2,1,2} l_{2,2,2} \{r_3 l_{3,1,2} l_{3,2,2} + \\ (m_{3,2}/r_2)\} + m_{2,2}/r_2 \{u_2 r_3 l_{3,1,2} l_{3,2,2} + ((1-u_2) m_{3,2}/r_3)\} \\ l_{3,1,2} l_{3,2,2} + (m_{3,2}/r_3) \end{bmatrix},$$

where the l_{t,j,k_t} and m_{t,k_t} notation remains as before with parameter set $\boldsymbol{\theta}_{sru} = [s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}, s_{3,1}, s_{3,2}, r_1, r_2, r_3, u_1, u_2]^T$, and

$$\begin{aligned} c_{10} \;=\; & r_1 l_{1,1,2} l_{1,2,2} \left(r_2 l_{2,1,2} l_{2,2,2} \left[r_3 l_{3,1,2} l_{3,2,2} + \{m_{3,2}/r_3\}\right] + \right. \\ & m_{2,2}/r_2 \left[u_2 r_3 l_{3,1,2} l_{3,2,2} + \{(1-u_2) m_{3,2}/r_3\}\right]) \\ & + m_{1,2}/r_1 \left(u_1 r_2 l_{2,1,2} l_{2,2,2} \left[r_3 l_{3,1,2} l_{3,2,2} + \{m_{3,2}/r_3\}\right]\right) \\ & \left. + (1-u_1) m_{2,2}/r_2 \left[u_2 r_3 l_{3,1,2} l_{3,2,2} + \{(1-u_2) m_{3,2}/r_3\}\right]\right). \end{aligned}$$

Consider this exhaustive summary as the first stage of a two-stage extension theorem proof as

$$\boldsymbol{\kappa}_1 = \begin{bmatrix} v \\ w_{1,1} \\ w_{1,2} \\ w_{2,1} \\ w_{2,2} \\ w_{3,1} \\ w_{3,2} \\ x_{[2,2]} \\ x_{[2,3]} \\ x_{[3,3]} \end{bmatrix} = \begin{bmatrix} r_1 r_2 r_3 \\ r_1/\left(s_{1,1}^2 s_{1,2}\right) - 1 \\ r_1/\left(s_{1,1} s_{1,2}^2\right) - 1 \\ r_2/\left(s_{2,1}^2 s_{2,2}\right) - 1 \\ r_2/\left(s_{2,1} s_{2,2}^2\right) - 1 \\ r_3/\left(s_{3,1}^2 s_{3,2}\right) - 1 \\ r_3/\left(s_{3,1} s_{3,2}^2\right) - 1 \\ l_{2,1,2} l_{2,2,2} + (m_{2,2} u_2 / r_2) \\ r_2 l_{2,1,2} l_{2,2,2} \left\{ r_3 l_{3,1,2} l_{3,2,2} + \right. \\ \left. (m_{3,2}/r_2) \right\} + m_{2,2} u_2 / r_2 \left\{ r_3 l_{3,1,2} l_{3,2,2} + (m_{3,2}/r_3) \right\} \\ l_{3,1,2} l_{3,2,2} + (m_{3,2}/r_3) \end{bmatrix},$$

with parameters $\boldsymbol{\theta}_1 = [s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}, s_{3,1}, s_{3,2}, r_1, r_2, r_3, u_2]$. The derivative matrix $[\partial \boldsymbol{\kappa}_1 / \partial \boldsymbol{\theta}_1]$ has full rank 10. The second stage involves the rest of the exhaustive terms as $\boldsymbol{\kappa}_2 = [x_{[1,1]}, x_{[1,2]}, x_{[1,3]}]^T$, but only has the additional parameter u_1, so $\boldsymbol{\kappa}_2$ is trivially full rank. Therefore by the two-stage extension theorem of Theorem 3.1, the model can be extended in terms of the number of seasons in the study.

Therefore, **s**, **r** and **u** form an exhaustive summary for any dimension which completes part three.

Part four:

To complete the proof, we can use the original extension theorem of Theorem 2.2 to prove the final result given our simpler exhaustive summary found before. Consider the simplest case where there is only one season of two surveys. The exhaustive summary given for this model using the simpler exhaustive summary has the form

$$\boldsymbol{\kappa} = \begin{bmatrix} s_{1,1} \\ s_{1,2} \\ r_1 \end{bmatrix} = \begin{bmatrix} \varphi_0^{[1]} p_{1,1} \\ \varphi_0^{[1]} p_{1,2} \\ \varphi_0^{[1]} p_{1,1} p_{1,2} \end{bmatrix},$$

with parameters $\boldsymbol{\theta} = [p_{1,1}, p_{1,2}, \varphi_0^{[1]}]^T$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}\right] = \begin{bmatrix} \varphi_0^{[1]} & 0 & \varphi_0^{[1]} p_{1,2} \\ 0 & \varphi_0^{[1]} & \varphi_0^{[1]} p_{1,1} \\ p_{1,1} & p_{1,2} & p_{1,1} p_{1,2} \end{bmatrix},$$

has full rank 3. This means that the original model is full rank. We can extend in two directions: By adding another survey and by adding another season. The first extension is a trivial application of the extension theorem as while it adds the exhaustive summary term $\varphi_0^{[1]} p_{1,3}$ and changes the previous exhaustive summary term $\varphi_0^{[1]} p_{1,1} p_{1,2}$ to $\varphi_0^{[1]} p_{1,1} p_{1,2} p_{1,3}$, it only adds the additional parameter $p_{1,3}$ so the extension is trivially full rank. In the second extension case, if we add another season also with two surveys, then the extended exhaustive terms are

$$\boldsymbol{\kappa}_2 = \begin{bmatrix} s_{2,1} \\ s_{2,2} \\ r_2 \\ u_1 \end{bmatrix} = \begin{bmatrix} \varphi_1^{[1,1]} p_{2,1} \\ \varphi_1^{[1,1]} p_{2,2} \\ \varphi_1^{[1,1]} p_{2,1} p_{2,2} \\ \varphi_1^{[0,1]} / \varphi_1^{[1,1]} \end{bmatrix},$$

with new parameters $\boldsymbol{\theta}_2 = [p_{2,1}, p_{2,2}, \varphi_1^{[0,1]}, \varphi_1^{[1,1]}]^T$. This extension has the derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_2}{\partial \boldsymbol{\theta}_2}\right] = \begin{bmatrix} \varphi_1^{[1,1]} & 0 & \varphi_1^{[1,1]} p_{2,2} & 0 \\ 0 & \varphi_1^{[1,1]} & \varphi_1^{[1,1]} p_{2,1} & 0 \\ 0 & 0 & 0 & 1/\varphi_1^{[1,1]} \\ p_{2,1} & p_{2,2} & p_{2,1} p_{2,2} & -\varphi_1^{[0,1]} / \left(\varphi_1^{[1,1]}\right)^2 \end{bmatrix},$$

which has full rank 4. This then proves that the model is full rank for any dimension, given that each season has at least two surveys.

To prove that the deficiency is increased by one for every season which has either zero or one survey during the season, we can consider what happens to the exhaustive summary when a new season is added with either zero or one survey in it. If we use the simpler exhaustive summary then adding a season with only one survey adds the exhaustive summary terms

$$\boldsymbol{\kappa}_{ext,1} = \begin{bmatrix} s_{t+1,1} = r_{t+1} \\ u_t \end{bmatrix} = \begin{bmatrix} \varphi_t^{[1,1]} p_{t+1,1} \\ \varphi_t^{[0,1]} / \varphi_t^{[1,1]} \end{bmatrix},$$

for the new extended season t with new parameters $\boldsymbol{\theta}_1 = [\varphi_t^{[1,1]}, \varphi_t^{[0,1]}, p_{t+1,1}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_{ext,1}}{\partial \boldsymbol{\theta}_1}\right] = \begin{bmatrix} p_{t+1,1} & -\varphi_t^{[0,1]} / \left(\varphi_t^{[1,1]}\right)^2 \\ 0 & 1/\varphi_t^{[1,1]} \\ \varphi_t^{[1,1]} & 0 \end{bmatrix},$$

shows us that the extension has a rank of 2 and a parameter deficiency of 1. This then shows the parameter deficiency goes up by 1 if there is only one survey in a season. If the extended season has no surveys at all, then this adds the exhaustive summary terms

$$\boldsymbol{\kappa}_{ext,0} = \begin{bmatrix} u_t \end{bmatrix} = \begin{bmatrix} \varphi_t^{[0,1]} / \varphi_t^{[1,1]} \end{bmatrix},$$

for the new extended season t with new parameters $\boldsymbol{\theta}_0 = [\varphi_t^{[1,1]}, \varphi_t^{[0,1]}]$. The derivative matrix

$$\left[\frac{\partial \boldsymbol{\kappa}_{ext,0}}{\partial \boldsymbol{\theta}_0}\right] = \begin{bmatrix} -\varphi_t^{[0,1]} / \left(\varphi_t^{[1,1]}\right)^2 \\ 1/\varphi_t^{[1,1]} \end{bmatrix},$$

shows us that the extension has a rank of 1 and a parameter deficiency of 1. This then shows the parameter deficiency goes up by 1 if there are no surveys in a season as well. This then completes the proof of Theorem 5.1. □

We note that there are other alternative reparameterisations that could have been used in this proof, such as for example the proof for when all occupancy and detection parameters are constant using the alternative reparameterisation given in Morgan et al. (2007). An alternative exhaustive summary would obtain the same general intrinsic parameter redundancy result of Theorem 5.1 as shown by use of the reparameterisation theorem of Theorem 3.2.

Bibliography

Amstrup, S. C., McDonald, T. L., and Manly, B. F. J. (2005). *The Handbook of Capture–Recapture Analysis.* Princeton University Press, Woodstock, Oxfordshire, UK.

Anderson, D. R. and Burnham, K. P. (1980). Effect of delayed reporting of band recoveries on survival estimates. *Journal of Field Ornithology,* **51,** 244–247.

Anderson, D. R., Burnham, K. P., and White, G. C. (1985). Problems in estimating age–specific survival rates from recovery data of birds ringed as young. *Journal of Animal Ecology,* **54,** 89–98.

Anderson, D. R. and Sterling, R. T. (1974). Population dynamics of molting pintail drakes banded in south–central Saskatchewan. *Journal of Wildlife Management,* **38,** 266–274.

Arnason, A. N. (1972). Parameter estimation from mark–recapture experiments on two populations subject to migration and death. *Researches on Population Ecology,* **13,** 97–113.

Arnason, A. N. and Mills, K. H. (1981). Bias and loss of precision due to tag loss in Jolly-Seber estimates for mark-recapture experiments. *Canadian Journal of Fisheries and Aquatic Sciences,* **38,** 1077–1095.

Audoly, S., D'Angió, L., Saccomani, M. P., and Cobelli, C. (1998). Global identifiability of linear compartmental models - A computer algebra algorithm. *IEEE Transactions on Biomedical Engineering,* **45,** 36–47.

Azuma, D. L., Baldwin, J. A., and Noon, B. R. (1990). Estimating the occupancy of spotted owl habitat areas by sampling and adjusting bias. Technical report, USDA Forest Service General Report PSW–124.

Bailey, L. L., Converse, S. J., and Kendall, W. L. (2010). Bias, precision, and parameter

redundancy in complex multistate models with unobservable states. *Ecology,* **91,** 1598–1604.

Bailey, L. L., MacKenzie, D. I., and Nichols, J. D. (2014). Advances and applications of occupancy models. *Methods in Ecology and Evolution,* in press.

Bailey, L. L., Reid, J. A., Forsman, E. D., and Nichols, J. D. (2009). Modeling co–occurrence of northern spotted and barred owls: Accounting for detection probability differences. *Biological Conservation,* **142,** 2983–2989.

Bailey, L. L., Simons, T. R., and Pollock, K. H. (2004a). Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications,* **14,** 692–702.

Bailey, L. L., Simons, T. R., and Pollock, K. H. (2004b). Spatial and temporal variation in detection probability of Plethodon salamanders using the robust capture-recapture design. *The Journal of Wildlife Management,* **68,** 14–24.

Barker, R. J. (1995). *Open Population Mark–Recapture Models Including Ancillary Sightings*. PhD thesis, Massey University, Palmerston North, New Zealand.

Barker, R. J. (1997). Joint modeling of live–recapture, tag–resight and tag–recovery data. *Biometrics,* **53,** 666–677.

Beale, C. M., Baker, N. E., Brewer, M. J., and Lennon, J. J. (2012). Protected area networks and savannah bird biodiversity in the face of climate change and land degradation. *Ecology Letters,* **16,** 1061–1068.

Bekker, P. A., Merckens, A., and Wansbeek, T. J. (1994). *Identification, Equivalent Models, and Computer Algebra*. Academic Press, London, UK.

Bellman, R. and Åmström, K. J. (1970). On structural identifiability. *Mathematical Biosciences,* **7,** 329–339.

Bellu, G., Saccomani, M. P., Audoly, S., and D'Angió, L. (2007). DAISY: A new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine,* **88,** 52–61.

Berthier, F., Diard, J. P., Pronzato, L., and Walter, E. (1996). Identifiability and distinguishability concepts in electrochemistry. *Automatica,* **32,** 973–984.

Birt, B., Rodwell, L. D., and Richards, J. P. (2009). Investigation into the sustainability of organic aquaculture of Atlantic cod (*Gadus morhua*). *Sustainability: Science, Practice and Policy,* **5,** 4–14.

Borchers, D. (2012). A non–technical overview of spatially explicit capture-recapture models. *Journal of Ornithology (Supplement 2),* **152,** S435–S444.

Borysiewicz, R. S. (2008). *The Combined Analysis of Multi–Site Mark–Recapture-Recovery Data and Multi–Site Census Data.* PhD thesis, The University of Kent, Canterbury, Kent, UK.

Brooks, S. P., Catchpole, E. A., and Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science,* **15,** 357–376.

Brown, D. I. (2010). *Climate Modelling for Animal Survival.* PhD thesis, School of Mathematics, Statistics and Actuarial Science, The University of Kent, Canterbury, UK.

Brownie, C., Anderson, D. R., Burnham, K. P., and Robson, D. S. (1985). *Statistical Inference from Band Recovery Data – A Handbook.* U.S. Dept of Interior, Fish and Wildlife Service, Washington, USA.

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H., and Hestbeck, J. B. (1993). Capture–recapture studies for multiple strata including non–Markovian transitions. *Biometrics,* **49,** 1173–1187.

Buckland, S. T. (1980). A modified analysis of the Jolly–Seber capture–recapture model. *Biometrics,* **36,** 419–435.

Buckland, S. T. (1982). A mark–recapture survival analysis. *Journal of Animal Ecology,* **51,** 833–847.

Burnham, K. P. (1993). A theory for combined analysis of ring recovery and recapture data. In Lebreton, J. D. and North, P. M., editors, *Marked Individuals in the Study of Bird Population*, pages 199–213. Bärenreiter Publishing Group, Basel, Switzerland.

Catchpole, E. A., Freeman, S. N., and Morgan, B. J. T. (1996). Steps to parameter redundancy in age–dependent recovery models. *Journal of the Royal Statistical Society: Series B,* **58,** 763–774.

Catchpole, E. A., Freeman, S. N., Morgan, B. J. T., and Harris, M. P. (1998). Integrated recovery/recapture data analysis. *Biometrics,* **54,** 33–46.

Catchpole, E. A., Kgosi, P. M., and Morgan, B. J. T. (2001). On the near–singularity of models for animal recovery data. *Biometrics,* **57,** 720–726.

Catchpole, E. A. and Morgan, B. J. T. (1991). A note on Seber's model for ring–recovery data. *Biometrika,* **78,** 917–919.

Catchpole, E. A. and Morgan, B. J. T. (1994). Boundary estimation in ring recovery models. *Journal of Royal Statistical Society: Series B,* **56,** 385–391.

Catchpole, E. A. and Morgan, B. J. T. (1996). Model selection in ring–recovery models using score tests. *Biometrics,* **52,** 664–672.

Catchpole, E. A. and Morgan, B. J. T. (1997). Detecting parameter redundancy. *Biometrika,* **84,** 187–196.

Catchpole, E. A. and Morgan, B. J. T. (2001). Deficiency of parameter–redundant models. *Biometrika,* **88,** 593–598.

Catchpole, E. A., Morgan, B. J. T., Coulson, T. N., Freeman, S. N., and Albon, S. D. (2000). Factors influencing Soay sheep survival. *Journal of the Royal Statistical Society: Series C,* **49,** 453–472.

Catchpole, E. A., Morgan, B. J. T., and Freeman, S. N. (1998). Estimation in parameter–redundant models. *Biometrika,* **85,** 462–468.

Catchpole, E. A., Morgan, B. J. T., Freeman, S. N., and Peach, W. J. (1999). Modelling the survival of British Lapwing *Vanellus vanellus* using ring–recovery data and weather covariates. *Bird Study,* **46,** S5–13.

Catchpole, E. A., Morgan, B. J. T., and Viallefont, A. (2002). Solving problems in parameter redundancy using computer algebra. *Journal of Applied Statistics,* **29,** 625–636.

Chapman, D. G. (1951). Some properties of the hypergeometric distribution with applications to zoological sample censuses. *University of California Publication in Statistics,* **1,** 131–160.

Chapman, D. G. and Robson, D. S. (1960). The analysis of a catch curve. *Biometrics,* **16,** 354–368.

Chappell, M. J. and Gunn, R. N. (1998). A procedure for generating locally identifiable reparameterisations of unidentifiable non–linear systems by the similarity transformation approach. *Mathematical Biosciences,* **148,** 21–41.

Chis, O. T., Banga, J. R., and Balsa-Canto, E. (2011). Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE,* **6,** e27755.

Choquet, R. and Cole, D. J. (2012). A hybrid symbolic–numerical method for determining model structure. *Mathematical Biosciences,* **236,** 117–125.

Choquet, R., Reboulet, A. M., Pradel, R., Gimenez, O., and Lebreton, J. D. (2004). `M-SURGE`: New software specifically designed for multistate capture-recapture models. *Animal Biodiversity and Conservation,* **27,** 207–215.

Cole, D. J. (2012). Determining parameter redundancy of multi–state mark–recapture models of sea birds. *Journal of Ornithology (Supplementary 2),* **152,** S305–S315.

Cole, D. J. and Choquet, R. (2012). Parameter redundancy in models with individual random effects. Technical Report.

Cole, D. J. and Freeman, S. N. (2012). Estimating age–specific survival rates from historical ringing data: Comment. Technical Report.

Cole, D. J. and McCrea, R. S. (2012). Parameter redundancy in discrete state–space and integrated models. Technical Report.

Cole, D. J. and Morgan, B. J. T. (2010a). A note on age–dependent tag return models for estimating fishing mortality, natural mortality, and selectivity. *Journal of Agricultural, Biological and Environmental Statistics,* **15,** 431–434.

Cole, D. J. and Morgan, B. J. T. (2010b). Parameter redundancy with covariates. *Biometrika,* **97,** 1002–1005.

Cole, D. J., Morgan, B. J. T., Catchpole, E. A., and Hubbard, B. A. (2012). Parameter redundancy in mark–recovery models. *Biometrical Journal,* **54,** 507–523.

Cole, D. J., Morgan, B. J. T., McCrea, R. S., Pradel, R., Gimenez, O., and Choquet, R. (2014). Does your species have memory? Analyzing capture–recapture data with memory models. *Ecology and Evolution,* **4,** 2124–2133.

Cole, D. J., Morgan, B. J. T., and Titterington, D. M. (2010). Determining the parametric structure of models. *Mathematical Biosciences,* **228,** 16–30.

Collier, B. A., Kremer, S. R., Mason, C. D., and Peterson, M. J. (2012). Survival, fidelity, and recovery rates of white–winged doves in Texas. *Journal of Wildlife Management,* **76,** 1129–1134.

Conroy, M. J. and Williams, B. K. (1984). A general methodology for maximum likelihood inference from band–recovery data. *Biometrics,* **40,** 739–748.

Cooch, E. G., Pradel, R., and Nur, N. (1997). *A Practical Guide to Mark–Recapture Analysis using* `SURGE` *(Second Edition).* Centre d´Ecologie Fonctionelle et Evolutive, CNRS, Montpellier, France.

Cooch, E. G. and White, G. C. (2014). Program `MARK`: A gentle introduction. `http://www.phidot.org/software/mark/docs/book`.

Corless, R. M. and Jeffrey, D. J. (1997). The Turing factorization of a rectangular matrix. *SIGSAM Bulletin,* **31,** 20–28.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika,* **51,** 429–438.

Cowen, L. and Schwarz, C. J. (2006). The Jolly-Seber model with tag loss. *Biometrics* **62,** 699–705.

Dochain, D., Vanrolleghem, P. A., and Van Daele, M. (1995). Structural identifiability of biokinetic models of activated sludge respiration. *Water Research,* **29,** 2571–2578.

Doligez, B., Clobert, J., Pettifor, R. A., Rowcliffe, M., Gustafsson, L., Perrins, C. M., and McCleery, R. H. (2002). Costs of reproduction: Assessing responses to brood size manipulation on life–history and behavioural traits using multi–state capture–recapture models. *Journal of Applied Statistics,* **29,** 407–423.

Dorazio, R. M. and Royle, J. A. (2005). Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association,* **100,** 389–398.

Dorazio, R. M., Royle, J. A., Söderström, B., and Glimskär, A. (2006). Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology,* **87,** 842–854.

Dreitz, V. J., Lukacs, P. M., and Knopf, F. L. (2006). Monitoring low density avian populations: An example using Mountain Plovers. *The Condor,* **108,** 700–706.

Eisenberg, M. C. and Hayashi, M. A. L. (2014). Determining structurally identifiable parameter combinations using subset profiling. To appear in Mathematical Biosciences.

Evans, N. D. and Chappell, M. J. (2000). Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems. *Mathematical Biosciences,* **168,** 137–159.

Falke, J. A., Fausch, K. D., Bestgen, K. R., and Bailey, L. L. (2010). Spawning phenology and habitat use in a Great Plains, USA, stream fish assemblage: An occupancy estimation approach. *Canadian Journal of Fisheries and Aquatic Sciences,* **67,** 1942–1956.

Fisher, F. M. (1959). Generalization of the rank and order conditions for identifiability. *Econometrica,* **27,** 431–447.

Fisher, F. M. (1961). Identifiability criteria in nonlinear systems. *Econometrica,* **29,** 574–590.

Fisher, F. M. (1963). Uncorrelated disturbances and identifiability criteria. *International Economic Review,* **4,** 134–152.

Fiske, I. J. and Chandler, R. B. (2011). `Unmarked`: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software,* **43,** 1–23.

Fiske, I. J., Chandler, R. B., Miller, D. A., Royle, J. A., and Kéry, M. (2014). `Unmarked`: Models for data from unmarked animals. `http://cran.r-project.org/web/packages/unmarked/index.html`.

Freeman, S. N. and Morgan, B. J. T. (1990). Studies in the analysis of ring–recovery data. *The Ring,* **13,** 217–287.

Freeman, S. N. and Morgan, B. J. T. (1992). A modelling strategy for recovery data from birds ringed as nestlings. *Biometrics,* **48,** 217–235.

Gardalia, T., Barton, D. C., White, J. D., and Geupel, G. R. (2003). Juvenile and adult survival of Swainson's thrush (*Catharus ustulatus*) in coastal California: Annual estimates using capture–recapture analyses. *The Auk,* **120,** 1188–1194.

Garrett, E. S. and Zeger, S. L. (2000). Latent class model diagnosis. *Biometrics,* **56,** 1055–1067.

Geissler, P. H. and Fuller, M. R. (1987). Estimation of the proportion of area occupied by an animal species. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association 1986*, pages 533–538.

Gelfand, A. E. and Sahu, S. K. (1999). Identifiability, improper priors, and Gibbs sampling for generalized linear models. *Journal of the American Statistical Association,* **94,** 247–253.

Gimenez, O., Blanc, L., Besnard, A., Pradel, R., Doherty Jr., P. F., Marboutin, E., and Choquet, R. (2014). Fitting occupancy models with `E–SURGE`: Hidden Markov modelling of presence–absence data. *Methods in Ecology and Evolution,* **5,** 592–597.

Gimenez, O., Choquet, R., and Lebreton, J. D. (2003). Parameter redundancy in multistate capture–recapture models. *Biometrical Journal,* **45,** 704–722.

Gimenez, O., Morgan, B. J. T., and Brooks, S. P. (2009). Weak identifiability in models for mark–recapture–recovery data. In Thomson, D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling Demographic Processes in Marked Populations*, volume 3 of *Springer Series: Environmental and Ecological Statistics*, pages 1055–1067. Springer, Dunedin, New Zealand.

Gimenez, O., Viallefont, A., Catchpole, E. A., Choquet, R., and Morgan, B. J. T. (2004). Methods for investigating parameter redundancy. *Animal Biodiversity and Conservation,* **27,** 561–572.

Godfrey, K. R., Jones, R. P., Browan, R. F., and Norton, J. P. (1982). Factors affecting the identifiability of compartmental models. *Automatica,* **18,** 285–293.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika,* **61,** 215–231.

Gould, W. R., Patla, D. A., Daley, R., Corn, P. S., Hossack, B. R., Bennetts, R., and Peterson, C. R. (2012). Estimating occupancy in large landscapes: Evaluation of amphibian monitoring in the Greater Yellowstone ecosystem. *Wetlands,* **32,** 379–389.

Green, A. W., Bailey, L. L., and Nichols, J. D. (2011). Exploring sensitivity of a multistate occupancy model to inform management decisions. *Journal of Applied Ecology,* **48,** 1007–1016.

Guillera-Arroita, G. (2011). Impact of sampling with replacement in occupancy studies with spatial replication. *Methods in Ecology and Evolution,* **2,** 401–406.

Guillera-Arroita, G., Morgan, B. J. T., Ridout, M. S., and Linkie, M. (2011). Species occupancy modeling for detection data collected along a transect. *Journal of Agricultural, Biological, and Environmental Statistics,* **16,** 301–317.

Guillera-Arroita, G., Ridout, M. S., Morgan, B. J. T., and Linkie, M. (2012). Models for species–detection data collected along transects in the presence of abundance–induced heterogeneity and clustering in the detection process. *Methods in Ecology and Evolution,* **3,** 358–367.

Haldane, J. B. S. (1955). The calcuation of mortality rates from ringing data. *Proceedings of the International Ornithological Congress,* **9,** 454–458.

Hènaux, V., Bregnballe, T., and Lebreton, J. D. (2007). Dispersal and recruitment during population growth in a colonial bird, the great cormorant *Phalacrocorax carbo sinensis*. *Journal of Avian Biology,* **38,** 44–57.

Hines, J. E., Nichols, J. D., Royle, J. A., MacKenzie, D. I., Gopalaswamy, A. M., Kumar, N. S., and Karanth, K. U. (2010). Tigers on trails: Occupancy modeling for cluster sampling. *Ecological Applications,* **20,** 14561466.

Hubbard, B. A., Cole, D. J., and Morgan, B. J. T. (2014). Parameter redundancy in capture–recapture–recovery models. *Statistical Methodology,* **17,** 17–29.

Hubbard, B. A., Cole, D. J., and Morgan, B. J. T. (2015). Parameter redundancy in occupancy models. Work in progress.

Hunter, C. M. and Caswell, H. (2009). Rank and redundancy of multistate mark–recapture models for seabird populations with unobservable states. *Modelling Demographic Processes in Marked Populations; Series: Environmental and Ecological Statistics,* **3,** 797–826.

Jiang, H., Pollock, K. H., Browne, C., Hightower, J. E., Hoenig, J. M., and Hearn, W. S. (2007). Age–dependent tag return models for estimating fishing mortality, natural mortality, and selectivity. *Journal of Agricultural, Biological and Environmental Statistics,* **12,** 177–194.

Jolly, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration–stochastic model. *Biometrika,* **52,** 225–247.

Kendall, W. L., Conn, P. B., and Hines, J. E. (2006). Combining multistate capture–recapture data with tag recoveries to estimate demographic parameters. *Ecology,* **87,** 169–177.

Kéry, M., Dorazio, R. M., Soldaat, L., van Strien, A., Zuiderwijk, A., and Royle, J. A. (2009). Trend estimation in populations with imperfect detection. *Journal of Applied Ecology,* **46,** 1163–1172.

King, R. (2012). A review of Bayesian state–space modelling of capture-recapture-recovery data. *Interface Focus,* **2,** 190–204.

King, R. (2014). Statistical ecology. *Annual Review of Statistics and Its Application,* **1,** 401–426.

King, R. and Brooks, S. P. (2003). Closed–form likelihoods for Arnason–Schwarz models. *Biometrika,* **90,** 435–444.

Koopmans, T. C., Rubin, H., and Leipnik, R. B. (1950). Measuring the equation systems of dynamic economics. In Koopmans, T. C. and Marschak, J., editors, *Statistical Inference in Dynamic Economic Models (Cowles Commission Monograph 10)*. Chapman and Hall, London, UK.

Krementz, D. G., Barker, R. J., and Nichols, J. D. (1997). Sources of variation in waterfowl survival rates. *The Auk,* **114,** 93–102.

Lakhani, K. H. and Newton, I. (1983). Estimating age–specific bird survival rates from ring recoveries – can it be done? *Journal of Animal Ecology,* **52,** 83–91.

Langrock, R. and King, R. (2013). Maximum likelihood estimation of mark–recapture–recovery models in the presence of continuous covariates. *Annals of Applied Statistics,* **7,** 1709–1732.

Laplace, P. S. (1786). Sur les naissances, les mariages et les morts. *Mèmoires de l' Acadèmie Royale des Sciences de Paris* .

Lebreton, J. D., Burnham, K. P., Clobert, J., and Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs,* **62,** 67–118.

Lebreton, J. D., Morgan, B. J. T., Pradel, R., and Freeman, S. N. (1995). A simultaneous analysis of dead recovery and live recapture data. *Biometrics,* **51,** 1418–1428.

Lincoln, F. C. (1930). *Calculating Waterfowl Abundance on the Basis of Banding Returns*. Number 118. U.S. Department of Agriculture, Washington D.C.

Lindberg, M. S. (2012). A review of designs for capture-mark-recapture studies in discrete time. *Journal of Ornithology (Supplement 2),* **152,** S355–S370.

Link, W. A. and Sauer, J. R. (1998). Estimating population change from count data: Application to the North American Breeding Bird Survey. *Ecological Applications,* **8,** 258–268.

Link, W. A., Yoshizaki, J., Bailey, L. L., and Pollock, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics,* **66,** 178–185.

MacKenzie, D. I., Bailey, L. L., Hines, J. E., and Nichols, J. D. (2011). An integrated model of habitat and species occurrence dynamics. *Methods in Ecology and Evolution,* **2,** 612–622.

MacKenzie, D. I., Bailey, L. L., and Nichols, J. D. (2004). Investigating species co–occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology,* **73,** 546–555.

MacKenzie, D. I. and Hines, J. (2014). `PRESENCE` 6.2: User manual. `http://www.mbr-pwrc.usgs.gov/software/doc/presence/presence.html`.

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., and Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology,* **84,** 2200–2207.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology,* **83,** 2248–2255.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., and Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Academic Press, USA.

MacKenzie, D. I., Nichols, J. D., Seamans, M. E., and Gutiérrez, R. J. (2009). Modeling species occurrence dynamics with multiple states and imperfect detection. *Ecology,* **90,** 823–835.

MacKenzie, D. I. and Royle, J. A. (2005). Designing occupancy studies: General advice and allocating survey effort. *Journal of Applied Ecology,* **42,** 1105–1114.

Mardekian, S. Z. and McDonald, L. (1981). Simultaneous analysis of band recovery and live–recapture data. *Journal of Wildlife Management,* **45,** 484–488.

Marzolin, G. (1988). Polygynie du cincle plongeur (*Cinclus cinclus*) dans les côtes de lorraine. *L´Oiseau et la Revue Francaise d´Ornithologie,* **58,** 277–286.

Matechou, E. (2010). *Applications and Extensions of Capture–Recapture Stop–Over Models*. PhD thesis, School of Mathematics, Statistics and Actuarial Science, The University of Kent, Canterbury, UK.

Matechou, E., Morgan, B. J. T., Pledger, S., Collazo, J. A., and Lyons, J. E. (2013). Integrated analysis of capture-recapture-resighting data and counts of unmarked

birds at stop–over sites. *Journal of Agricultural, Biological, and Environmental Statistics,* **18,** 120–135.

Mazerolle, M. J., Bailey, L. L., Kendall, W. L., Royle, J. A., Converse, S. J., and Nichols, J. D. (2007). Making great leaps forward: Accounting for detectability in herpetological field studies. *Journal of Herpetology,* **41,** 672–689.

McCrea, R. S. (2012). Sufficient statistic likelihood construction for age– and time–dependent multi–state joint recapture and recovery data. *Statistics and Probability Letters,* **82,** 357–359.

McCrea, R. S. and Morgan, B. J. T. (2008). A history of capture–recapture in ecology. `http://www.ncse.org.uk/files/2014/04/CapDataHistory.pdf`.

McCrea, R. S., Morgan, B. J. T., and Bregnballe, T. (2012). Model comparison and assessment for multi–state capture–recapture–recovery models. *Journal of Ornithology (Supplement 2),* **152,** S293–S303.

McCrea, R. S., Morgan, B. J. T., and Cole, D. J. (2013). Age–dependent mixture models for recovery data on animals marked at unknown age. *Journal of the Royal Statistical Society: Series C,* **62,** 101–113.

McCrea, R. S., Morgan, B. J. T., Gimenez, O., Besbeas, P. T., Lebreton, J. D., and Bregnballe, T. (2010). Multi–site integrated population modelling. *Journal of Agricultural, Biological, and Environmental Statistics,* **15,** 539–561.

McCrea, R. S., Morgan, B. J. T., and Pradel, R. (2013). Diagnostic goodness–of–fit tests for joint recapture and recovery models. *The Journal of Agricultural, Biological, and Environmental Statistics,* in press.

McDonald, T. L., Amstrup, S. C., and Manly, B. F. J. (2003). Tag loss can bias Jolly–Seber capture-recapture estimates. *Wildlife Society Bulletin,* **31,** 814–822.

Mead, C. J., North, P. M., and Watmough, B. R. (1979). The mortality of British Grey Herons. *Bird Study,* **26,** 13–22.

Meyer, C. D. (1973). Generalized inverses and ranks of block matrices. *SIAM Journal on Applied Mathematics,* **25,** 597–602.

Miller, D. A., Nichols, J. D., McClintock, B. T., Campbell Grant, E. H., Bailey, L. L., and Weir, L. A. (2011). Improving occupancy estimation when two types of observational error occur: Non–detection and species misidentification. *Ecology,* **92,** 1422–1428.

Morgan, B. J. T. (2008). *Applied Stochastic Modelling.* CRC Press, London, UK, Second edition.

Morgan, B. J. T. and Freeman, S. N. (1989). A model with first–year variation for ring–recovery data. *Biometrics,* **45,** 1087–1101.

Morgan, B. J. T., Revell, D. J., and Freeman, S. N. (2007). A note on simplifying likelihoods for site occupancy models. *Biometrics,* **63,** 618–621.

Nams, V. O. and Gillis, E. A. (2003). Changes in tracking tube use by small mammals over time. *Journal of Mammalogy,* **84,** 1374–1380.

Nasution, M. D., Browne, C., Pollock, K. H., and Powell, R. A. (2004). The effect of model identifiability of allowing different relocation rates for live and dead animals in the combined analysis of telemetry and recapture data. *Journal of Agricultural, Biological and Environmental Statistics,* **9,** 27–41.

Nichols, J. D., Bailey, L. L., O'Connell Jr., A. F., Talancy, N. W., Campbell Grant, E. H., Gilbert, A. T., Annand, E. M., Husband, T. P., and Hines, J. E. (2008). Multi–scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology,* **45,** 1321–1329.

Nichols, J. D., Hines, J. E., MacKenzie, D. I., Seamans, M. E., and Gutiérrez, R. J. (2007). Occupancy estimation and modeling with multiple states and state uncertainty. *Ecology,* **88,** 1395–1400.

Nichols, J. D. and Kendall, W. L. (1995). The use of multi–state capture–recapture models to address questions in evolutionary ecology. *Journal of Applied Statistics,* **22,** 835–846.

Nichols, J. D., Stokes, S. L., Hines, J. E., and Conroy, M. J. (1982). Additional comments on the assumption of homogeneous survival rates in modern bird banding estimation models. *The Journal of Wildlife Management,* **46,** 953–962.

North, P. M. and Morgan, B. J. T. (1979). Modelling heron survival using weather data. *Biometrics,* **35,** 667–681.

O'Connell Jr., A. F., Talancy, N. W., Bailey, L. L., Sauer, J. R., Cook, R., and Gilbert, A. T. (2006). Estimating site occupancy and detection probability parameters for meso– and large mammals in a coastal ecosystem. *The Journal of Wildlife Management,* **70,** 1625–1633.

Otto, C. R. V., Bailey, L. L., and Roloff, G. J. (2013). Improving species occupancy estimation when sampling violates the closure assumption. *Ecography,* **36,** 1299–1309.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., and Peterson, A. T. (2007). Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography,* **34,** 102–117.

Petersen, C. G. J. (1896). The yearly immigration of young plaice into the Limfiord from the German Sea. *Report of the Danish Biological Station (1895),* **6,** 5–84.

Pledger, S., Efford, M. G., Pollock, K. H., Collazo, J., and Lyons, J. (2009). Stopover duration analysis with departure probability dependent on unknown time since arrival. In Thomson, D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling Demographic Processes in Marked Populations: Environmental and Ecological Statistics Volume 3*, pages 349–363. Springer, New York, USA.

Pledger, S., Pollock, K. H., and Norris, J. L. (2003). Open capture–recapture models with heterogeneity: I. Cormack–Jolly–Seber model. *Biometrics,* **59,** 786–794.

Pollock, K. H. (1982). A capture–recapture design robust to unequal probability of capture. *The Journal of Wildlife Management,* **46,** 752–757.

Pollock, K. H. (2000). Capture–recapture models. *Journal of the American Statistical Association,* **95,** 293–296.

Pollock, K. H., Nichols, J. D., Brownie, C., and Hines, J. E. (1990). Statistical inference for capture–recapture experiments. *Wildlife Monographs,* **107,** 3–97.

Pollock, K. H. and Raveling, D. G. (1982). Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. *The Journal of Wildlife Management,* **46,** 88–98.

Pomeroy, P., Smout, S., Moss, S., Twiss, S., and King, R. (2010). Low and delayed recruitment at two grey seal breeding colonies in the UK. *Journal of Northwest Atlantic Fisheries Science,* **42,** 125–133.

Pradel, R. (2005). Multievent: An extension of multistate capture–recapture models to uncertain states. *Biometrics,* **61,** 442–447.

Pradel, R. and Sanz-Aguilar, A. (2012). Modeling trap–awareness and related phenomena in capture–recapture studies. *PLoS ONE,* **7,** 1–4.

Reiersøl, O. (1950). Identifiability of a linear relation between variables which are subject to error. *Econometrica,* **18,** 375–389.

Richmond, O. M. W., Hines, J. E., and Beissinger, S. R. (2010). Two–species occupancy models: A new parameterization applied to co–occurrence of secretive rails. *Ecological Applications,* **20,** 2036–2046.

Ridout, M., Hinde, J., and Demétrio, C. G. B. (2001). A score test for testing a zero–inflated Poisson regression model against zero–inflated negative binomial alternatives. *Biometrics,* **57,** 219–223.

Robbins, C. S., Bystrak, D., and Geissler, P. H. (1986). *The Breeding Bird Survey: Its First Fifteen Years, 1965–1979.* Number 157. Publication Series of the U.S. Fish and Wildlife Service Resource Publication.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica,* **39,** 577–591.

Royle, J. A. (2004a). Modeling abundance index data from anuran calling surveys. *Conservation Biology,* **18,** 1378–1385.

Royle, J. A. (2004b). N–mixture models for estimating population size from spatially replicated counts. *Biometrics,* **60,** 108–115.

Royle, J. A. (2006). Site occupancy models with heterogeneous detection probabilities. *Biometrics,* **62,** 97–102.

Royle, J. A. (2008). Modeling individual effects in the Cormack–Jolly–Seber model: A state-space formulation. *Biometrics,* **64,** 364–370.

Royle, J. A. and Link, W. A. (2005). A general class of multinomial mixture models for Anuran calling survey data. *Ecology,* **86,** 2505–2512.

Royle, J. A. and Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology,* **87,** 835–841.

Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology,* **84,** 777–790.

Saraux, C., Behoc, C. L., Durant, J. M., Viblanc, V. A., Gauthier-Clerc, M., Beaune, D., Park, Y. H., Yoccoz, N. G., Stenseth, N. C., and Maho, Y. L. (2011). Reliability of flipper–banded penguins as indicators of climate change. *Nature,* **469,** 203–206.

Sauer, J. R., Peterjohn, B. G., and Link, W. A. (1994). Observer differences in the North American Breeding Bird Survey. *The Auk,* **111,** 50–62.

Schnabel, Z. E. (1938). The estimation of total fish population of a lake. *The American Mathematical Monthly,* **45,** 348–352.

Schwarz, C. J. and Stobo, W. T. (2000). Estimation of juvenile survival, adult survival, and age–specific pupping probabilities for the female grey seal (*Halichoerus gryprus*) on Sable Island from capture–recapture data. *Canadian Journal of Fisheries and Aquatic Sciences,* **57,** 247–253.

Seber, G. A. F. (1965). A note on the multiple–recapture census. *Biometrika,* **52,** 249–259.

Seber, G. A. F. (1970). Estimating time–specific survival and reporting rates for adult birds from band returns. *Biometrika,* **57,** 313–318.

Seber, G. A. F. (1971). Estimating age–specific survival rates from bird–band returns when the reporting rate is constant. *Biometrika,* **58,** 491–497.

Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics,* **42,** 267–292.

Seber, G. A. F. (2002). *The Estimation of Animal Abundance and Related Parameters.* Blackburn Press, Blackburn, Lancashire, UK.

Shapiro, A. (1986). Asymptotic theory of overparameterized structural models. *Journal of the American Statistical Association,* **81,** 142–149.

Shapiro, A. and Browne, M. W. (1983). On the investigation of local identifiability: A counterexample. *Psychometrika,* **48,** 303–304.

Silvey, S. D. (1975). *Statistical Inference.* Chapman and Hall, London, UK.

Stanley, T. R. and Royle, J. A. (2005). Estimating site occupancy and abundance using indirect detection indices. *The Journal of Wildlife Management,* **69,** 874–883.

Steen, D. A., McClure, C. J. W., Brock, J. C., Rudolph, D. C., Pierce, J. B., Lee, J. R., Humphries, W. J., Gregory, B. B., Sutton, W. B., Smith, L. L., Baxley, D. L., Stevenson, D. J., and Guyer, C. (2014). Snake co–occurrence patterns are best explained by habitat and hypothesized effects of interspecific interactions. *Journal of Animal Ecology,* **83,** 286–295.

Szymczak, M. and Rexstad, E. (1991). Harvest distribution and survival of a gadwall population. *Journal of Wildlife Management,* **55,** 592–600.

Thomas, L., Buckland, S. T., Rexstad, E. A., Laake, J. L., Strindberg, S., Hedley, S. L., Bishop, J. R. B., Marques, T. A., and Burnham, K. P. (2010). Distance software: Design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology,* **47,** 5–14.

Tilling, K. and Sterne, J. A. C. (1999). Capture–recapture models including covariate effects. *American Journal of Epidemiology,* **149,** 392–400.

Turing, A. M. (1948). Rounding–off errors in matrix processes. *Quarterly Journal of Mechanics and Applied Mathematics,* **1,** 287–308.

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K., and Possingham, H. P. (2003). Improving precision and reducing bias in biological surveys: Estimating false–negative error rates. *Ecological Applications,* **13,** 1790–1801.

van Wieringen, W. N. (2005). On identifiability of certain latent class models. *Statistics and Probability Letters,* **75,** 211–218.

Viallefont, A., Lebreton, J. D., and Reboulet, A. M. (1998). Parameter identifiability and model selection in capture–recapture models: A numerical approach. *Biometrical Journal,* **3,** 313–325.

Walter, E. and Lecourtier, Y. (1982). Global approaches to identifiability testing for linear and nonlinear state space models. *Mathematics and Computers in Simulation,* **24,** 472–482.

Walter, E. and Pronzato, L. (1996). On the identifiability and distinguishability of nonlinear parametric models. *Mathematics and Computers in Simulation,* **42,** 125–134.

Webb, S. L., Gee, K. L., and Wang, G. (2010). Survival and fidelity of an enclosed white–tailed deer population using capture-recapture–reporting data. *Population Ecology,* **52,** 81–88.

Weir, L. A., Fiske, I. J., and Royle, J. A. (2009). Trends in anuran occupancy from northeastern states of the North American Amphibian Monitoring Program. *Herpetological Conservation and Biology,* **4,** 389–402.

Weir, L. A. and Mossman, M. J. (2005). North American Amphibian Monitoring Program. In Lannoo, M. J., editor, *Declining Amphibians: A United States Response*

to the Global Phenomenon, pages 307–313. University of California Press, Berkeley, California, USA.

Weir, L. A., Royle, J. A., Nanjappa, P., and Jung, R. E. (2005). Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. *Journal of Herpetology,* **39,** 627–639.

Whiley, M. and Titterington, D. M. (2002). Model identifiability in naive Bayesian networks. Technical report, Glasgow, UK.

White, G. C. (1983). Numerical estimation of survival rates from band–recovery and biotelemetry data. *The Journal of Wildlife Management,* **47,** 716–728.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press Inc, San Diego, California, USA.

Yamaura, Y., Royle, J. A., Kuboi, K., Tada, T., Ikeno, S., and Makino, S. (2011). Modelling community dynamics based on species–level abundance models from detection/nondetection data. *Journal of Applied Ecology,* **48,** 67–75.

Yu, C., Cole, D. J., and Morgan, B. J. T. (2014). Parameter redundancy of mixture models in capture recapture data. Work in progress.